



BEST AVAILABLE COPY

# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 145331

TO: Janet Epps-Ford  
Location: rem/2c05/2c18  
Art Unit: 1635  
Tuesday, March 01, 2005

Case Serial Number: 10/086062

From: Barb O'Bryen  
Location: Biotech-Chem Library  
Remsen 1A69  
Phone: 571-272-2518

barbara.obryen @uspto.gov

### Search Notes

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

On nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:32:36 ; Search time 729.221 Seconds

(without alignments)  
1661.198 Million cell updates/sec

Title: US-10-086-062-1

Perfect score: 25

Sequence: 1 ctggacccctctcgagatccgt 25

Scoring table: IDENTITY\_NUC

Gapext 1.0

Searched: 4708233 seqb, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_be: \*  
2: gb\_ntg: \*  
3: gb\_in: \*  
4: gb\_on: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pi: \*  
9: gb\_pr: \*  
10: gb\_rx: \*  
11: gb\_stb: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No. Score Query Length DB ID Description

1 25 100.0 25 AX358106 Sequence

2 25 100.0 813 6 AR097833 Sequence

3 25 100.0 994 6 AR097835 Sequence

4 25 100.0 1807 6 AR097836 Sequence

5 25 100.0 2182 8 AT342393 Zea diplo

6 25 100.0 2620 6 AR097837 Sequence

7 25 100.0 2807 6 AX134224 Sequence

8 25 100.0 2949 6 AX259248 Sequence

9 25 100.0 3069 6 AX202417 Sequence

10 25 100.0 3121 6 AX202416 Sequence

11 25 100.0 3433 6 AR097838 Sequence

12 25 100.0 3840 6 138571 Sequence

13 25 100.0 3840 6 AR365568 Sequence

14 25 100.0 3841 8 S94464 polyribonuc

15 25 100.0 3972 6 AX55352 Sequence

16 25 100.0 4190 8 ZMU29159 Zea mays c1

17 25 100.0 4341 6 AX259249 Sequence

18 25 100.0 5465 6 BD194459 Chimeric

19 25 100.0 5465 6 BD194459 Chimeric

20 25 100.0 5522 12 AY452753 Sequence

21 25 100.0 5622 6 AX03929 Sequence

22 25 100.0 6103 6 BD194458 Chimeric

23 25 100.0 6111 6 BD247866 Transgeni

24 25 100.0 6111 6 AR193435 Sequence

25 25 100.0 6111 6 AR430220 Sequence

26 25 100.0 6215 6 AX118823 Sequence

27 25 100.0 6217 6 AR209726 Sequence

28 25 100.0 6579 6 AX118824 Sequence

29 25 100.0 6667 6 AX118825 Sequence

30 25 100.0 6818 6 AX840287 Sequence

31 25 100.0 6898 6 BD137177 Requiliator

32 25 100.0 6898 6 AR209020 Sequence

33 25 100.0 6898 6 AR428089 Sequence

34 25 100.0 7018 6 AY452736 Reporter

35 25 100.0 7027 6 AX840286 Sequence

36 25 100.0 9172 6 AX259246 Sequence

37 25 100.0 9359 6 AX384394 Sequence

38 25 100.0 9359 6 AX473364 Sequence

39 25 100.0 11643 6 AX840289 Sequence

40 25 100.0 12817 6 AX356664 Sequence

41 25 100.0 12949 6 BD247866 Transgeni

42 25 100.0 12949 6 AR193437 Sequence

43 25 100.0 12949 6 AR430222 Sequence

44 25 100.0 12982 6 BD091582 Vector fo

45 25 100.0 12982 6 BD091582 Vector fo

#### ALIGNMENTS

RESULT 1  
AX358106 Locus AX358106 Definition Sequence 1 from Patent WO0194394. Version AX358106.1 GI:18674853

ACCESSION AX358106  
KEYWORDS Agrobacterium tumefaciens (Rhizobium radiobacter)  
ORGANISM Agrobacterium tumefaciens (Rhizobacter; Proteobacteria; Alpha-proteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium; Bacteria)

REFERENCE 1 Jilka, J.M., Hood, B.B. and Howard, J.A.  
AUTHORS Novel plant promoter sequences and methods of use for same  
TITLE Patent: WO 0194394-A 1 13-DEC-2001;  
JOURNAL Prodigene, Inc. (US)

FEATURES Location/Qualifiers  
1. .25 /organism="Agrobacterium tumefaciens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:358"

#### ORIGIN

Query Match 100.0%; Score 25; DB 6; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.12%; Mismatches 0; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTCGAGAGTCGCT 25

Db 1 CTGGACCCCTCTCGAGAGTCGCT 25

RESULT 2  
AR097833 Locus AR097833 Definition Sequence 13 from patent US 6072050. Version AR097833.1 GI:12806563  
KEYWORDS Unknown  
ORGANISM Unclassified.

REFERENCE	1 (bases 1 to 813)
AUTHORS	Bowen,B.A., Bruce,W.B., Lu,G., Sims,L.E. and Tagliani,L.A.
TITLE	Synthetic promoters
JOURNAL	Patent: US 6072050-A 13 06-JUN-2000;
FEATURES	Location/Qualifiers
Source	1. .813 /organism="unknown" /mol_type="unassigned DNA"
ORIGIN	
Query Match	100.0%; Score 25; DB 6; Length 813;
Best Local Similarity	100.0%; Pred. No. 0.11;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CTGGACCCCTCTCGAGAGTCGCT 25
Db	654 CTGGACCCCTCTCGAGAGTCGCT 678
RESULT 3	
AR097835	AR097835 Sequence 15 from patent US 6072050. DNA linear PAT 14-FEB-2001
DEFINITION	994 bp
LOCUS	
ACCESSION	AR097835
VERSION	AR097835.1 GI:12806565
KEYWORDS	Unknown.
SOURCE	Unclassified.
ORGANISM	
REFERENCE	1 (bases 1 to 994)
AUTHORS	Bowen,B.A., Bruce,W.B., Lu,G., Sims,L.E. and Tagliani,L.A.
TITLE	Synthetic promoters
JOURNAL	Patent: US 6072050-A 15 06-JUN-2000;
FEATURES	Location/Qualifiers
Source	1. .994 /organism="unknown" /mol_type="unassigned DNA"
ORIGIN	
Query Match	100.0%; Score 25; DB 6; Length 994;
Best Local Similarity	100.0%; Pred. No. 0.11;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CTGGACCCCTCTCGAGAGTCGCT 25
Db	654 CTGGACCCCTCTCGAGAGTCGCT 678
RESULT 4	
AR097836	AR097836 Sequence 16 from patent US 6072050. DNA linear PAT 14-FEB-2001
DEFINITION	1807 bp
LOCUS	
ACCESSION	AR097836
VERSION	AR097836.1 GI:12806566
KEYWORDS	Unknown.
SOURCE	Unclassified.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 1807)
AUTHORS	Bowen,B.A., Bruce,W.B., Lu,G., Sims,L.E. and Tagliani,L.A.
TITLE	Synthetic promoters
JOURNAL	Patent: US 6072050-A 16 06-JUN-2000;
FEATURES	Location/Qualifiers
Source	1. .1807 /organism="unknown" /mol_type="unassigned DNA"
ORIGIN	
Query Match	100.0%; Score 25; DB 6; Length 1807;
Best Local Similarity	100.0%; Pred. No. 0.11;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CTGGACCCCTCTCGAGAGTCGCT 25
Db	654 CTGGACCCCTCTCGAGAGTCGCT 678
RESULT 5	
AR097837	AR097837 Sequence 17 from patent US 6072050. DNA linear PAT 14-FEB-2001
DEFINITION	2182 bp
LOCUS	
ACCESSION	AR097837
VERSION	AR097837.1 GI:12806567
KEYWORDS	Unknown.
SOURCE	Unclassified.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 2620)
AUTHORS	Bowen,B.A., Bruce,W.B., Lu,G., Sims,L.E. and Tagliani,L.A.
TITLE	Synthetic promoters
JOURNAL	Patent: US 6072050-A 17 06-JUN-2000;
FEATURES	Location/Qualifiers
Source	1. .2620 /organism="unknown" /mol_type="unassigned DNA"
ORIGIN	

Query Match	100.0%; Score 25; DB 6; Length 2620;
Best Local Similarity	100.0%; Pred. No. 0.11; 0; Mismatches 0; Indels 0; Gaps 0;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CTGGACCCCTCTCGAGAGTCGGCT 25
Db	654 CTGGACCCCTCTCGAGAGTCGGCT 678
RESULT 7	
AX134224	AX134224
DEFINITION	Sequence 1 from Patent WO0132897.
LOCUS	AX134224
ACCESSION	AX134224.1 GI:14270967
VERSION	AX202417.1 GI:15392164
KEYWORDS	
SOURCE	
ORGANISM	synthetic construct
REFERENCE	other sequences; artificial sequences.
AUTHORS	1. Groenewald,J.H. and Botha,F.C.
TITLE	A high level, stable, constitutive promoter element for plants
JOURNAL	Patent: WO 0132897-A 1 10-MAY-2001;
FEATURES	South African Sugar Association (ZA) ; University of Stellenbosch, Institute for Plant Biotechnology (ZA)
Source	Location/Qualifiers
1.	2807 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="synthetic construct"
ORIGIN	
Query Match	100.0%; Score 25; DB 6; Length 2807;
Best Local Similarity	100.0%; Pred. No. 0.11; 0; Mismatches 0; Indels 0; Gaps 0;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CTGGACCCCTCTCGAGAGTCGGCT 25
Db	1501 CTGGACCCCTCTCGAGAGTCGGCT 1525
RESULT 8	
AX259248	AX259248
DEFINITION	Sequence 46 from Patent WO0173087.
LOCUS	AX259248
ACCESSION	AX259248.1 GI:16508494
VERSION	AX202416.1 GI:15392163
KEYWORDS	
SOURCE	
ORGANISM	synthetic construct
REFERENCE	other sequences; artificial sequences.
AUTHORS	1. Barbas,C.F., Stege,J.T., Guan,X. and Dalmaia,B.
TITLE	Methods and compositions to modulate expression in plants
JOURNAL	Patent: WO 0152620-A 5 26-JUL-2001; The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL DISCOVERY, INC. (CA)
FEATURES	Location/Qualifiers
Source	
1.	3121 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="PND3018 coding region"
ORIGIN	
Query Match	100.0%; Score 25; DB 6; Length 3059;
Best Local Similarity	100.0%; Pred. No. 0.11; 0; Mismatches 0; Indels 0; Gaps 0;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CTGGACCCCTCTCGAGAGTCGGCT 25
Db	677 CTGGACCCCTCTCGAGAGTCGGCT 701
RESULT 10	
AX202416	AX202416
DEFINITION	Sequence 4 from Patent WO0152620.
LOCUS	AX202416
ACCESSION	AX202416
VERSION	AX202416.1 GI:15392163
KEYWORDS	
SOURCE	
ORGANISM	synthetic construct
REFERENCE	other sequences; artificial sequences.
AUTHORS	1. Barbas,C.F., Stege,J.T., Guan,X. and Dalmaia,B.
TITLE	Methods and compositions to modulate expression in plants
JOURNAL	Patent: WO 0152620-A 5 26-JUL-2001; The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL DISCOVERY, INC. (CA)
FEATURES	Location/Qualifiers
Source	
1.	3121 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="PND3008 coding region"
ORIGIN	
Query Match	100.0%; Score 25; DB 6; Length 3121;
Best Local Similarity	100.0%; Pred. No. 0.11; 0; Mismatches 0; Indels 0; Gaps 0;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CTGGACCCCTCTCGAGAGTCGGCT 25
Db	677 CTGGACCCCTCTCGAGAGTCGGCT 701
RESULT 11	
AR097838	AR097838
DEFINITION	Sequence 18 from patent US 6072050.
LOCUS	AR097838
ACCESSION	AR097838

VERSION AR097838.1 GI:12806568  
 KEYWORDS Unknown.  
 ORGANISM Unknown.

REFERENCE  
 AUTHORS Bowen, B.A., Bruce, W.B., Lu, G., Sims, L.E. and Tagliani, L.A.  
 TITLE Synthetic promoters  
 JOURNAL Patent: US 6720505 A 18 06-JUN-2000;  
 FEATURES SOURCE Location/Qualifiers  
 1. .3433  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

RESULT 12  
 LOCUS I06936 3840 bp DNA linear PAT 02-DRC-1994  
 DEFINITION Sequence 8 from Patent EP 0342926.  
 ACCESSION I06936  
 VERSION I06936.1 GI:589916  
 KEYWORDS SOURCE Unknown.  
 ORGANISM Unclassified.  
 UNCLASSIFIED 1 (bases 1 to 3840)  
 AUTHORS Quail, P.H., Christiansen, A.H., Hershey, H.P., Sharrock, R.A. and Sullivan, T.D.  
 TITLE JOURNAL Plant ubiquitin promoter system  
 ACCESSION EP 0342926-A2 8 23-NOV-1989;  
 VERSION FEATURES SOURCE 1. .3840  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Query Match 100.0%; Score 25; DB 6; Length 3843;  
 Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTCGAGAGTCGGT 25  
 Db 654 CTGGACCCCTCTCGAGAGTCGGT 678

RESULT 14  
 LOCUS AR365568 3840 bp DNA linear PAT 03-SEP-2003  
 DEFINITION Sequence 1 from patent US 5510474.  
 ACCESSION AR365568  
 VERSION AR365568.1 GI:34429351  
 KEYWORDS SOURCE Unknown.  
 ORGANISM Unclassified.  
 UNCLASSIFIED 1 (bases 1 to 3840)  
 AUTHORS Quail, P.H., Christiansen, A.H., Hershey, H.P., Sharrock, R.A. and Sullivan, T.D.  
 TITLE JOURNAL Plant ubiquitin promoter system  
 ACCESSION Patent: US 5510474-A 1 23-SEP-1996;  
 VERSION FEATURES SOURCE 1. .3840  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN  
 Query Match 100.0%; Score 25; DB 6; Length 3840;  
 Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTCGAGAGTCGGT 25  
 Db 686 CTGGACCCCTCTCGAGAGTCGGT 710

RESULT 15  
 LOCUS S94464 3841 bp DNA linear PLN 07-MAY-1993  
 DEFINITION polyubiquitin [maize, Genomic, 3841 nt].  
 ACCESSION S94464  
 VERSION S94464.1 GI:248336  
 KEYWORDS SOURCE Zea mays  
 ORGANISM Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidea; Andropogoneae; Zea.  
 REFERENCE  
 AUTHORS Christensen, A.H., Sharrock, R.A. and Quail, P.H.  
 TITLE Maize polyubiquitin genes, structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation  
 ACCESSION 1 (bases 1 to 3841)  
 VERSION JOURNAL Plant Mol. Biol. 18 (4), 675-689 (1992)  
 FEATURES MEDLINE 9216044  
 PUBLISHED 1313711  
 REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 94464] from the original journal article.  
 FEATURES SOURCE 1. .3841  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4577"  
 /organism="polyubiquitin"  
 /gene="polyubiquitin"  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 /codon\_start=1

RESULT 13  
 LOCUS I38571 3840 bp DNA linear PAT 13-MAY-1997  
 DEFINITION Sequence 1 from patent US 5614399.  
 ACCESSION I38571  
 VERSION I38571.1 GI:2084625  
 KEYWORDS SOURCE Unknown.  
 ORGANISM Unclassified.  
 UNCLASSIFIED 1 (bases 1 to 3840)  
 AUTHORS Quail, P.H., Christensen, A.H., Hershey, H.P., Sharrock, R.A. and Sullivan, T.D.  
 TITLE Plant ubiquitin promoter system  
 JOURNAL Patent: US 5614399-A 1 25-MAR-1997;  
 FEATURES LOCATION/QUALIFIERS 1. .3840  
 SOURCE /organism="unknown"  
 /mol\_type="unassigned DNA"

/protein\_id="AAB21993.1"  
/db\_xref="GI:248337"  
/translation="MQLFVKUTLGKMITLEVSSDTIDNYAKIQLQKEGIPPDQORLI  
FAGKQLEDEGRTIADNYIQLKESTHLVILRGGMQIFVKUTLGKMITLEVSSDTIDNV  
KAKIQDEKEGIPPDQORLIQAGKQLEGRTIADNYIQLKESTHLVILRGGMQIFVKUTL  
TGKTITLEVSSDTIDNYAKIQLDEGIPPDQORLIQAGKQLEGRTIADNYIQLKEST  
LHLVLRGGMQIFVKUTLGKMITLEVSSDTIDNYAKIQLKEGIPPDQORLIQAGK  
QDEGRTIADNYIQLKESTHLVILRGGMQIFVKUTLGKMITLEVSSDTIDNYAKI  
ITLEVSSDTIDNYAKIQLKEGIPPDQORLIQAGKQLEGRTIADNYIQLKESTHLV  
LRLRGCMQIFVKUTLGKMITLEVSSDTIDNYAKIQLKEGIPPDQORLIQAGKQLE  
GRTIADNYIQLKESTHLVILRGGQ"

ORIGIN

Query Match 100.0%; Score 25; DB 8; Length 3841;  
Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CTGGACCCCTCTCGAGATTCGCT	25
Db	687	CTGGACCCCTCTCGAGATTCGCT	711

Search completed: March 1, 2005, 00:29:10  
Job time : 735.221 sec8

This page blank (uspto)

OM nucleic - nucleic search, using SW model  
Run On: February 28, 2005 ; 23:26:41 ; Search time 218.182 Seconds  
Copyright (C) 1993 - 2005 Compugen Ltd.

ALIGNMENT

The invention relates to a promoter sequence capable of directing expression of a nucleotide sequence in a plant cell, comprising maize ubiquitin (Ubi-1) promoter sequence with a modification so that it does not include two overlapping heat shock elements (HSE) or it directs expression to increase the endosperm/embryo expression ratio of the protein when compared to the ratio from a wild-type ubiquitin promoter.

The modified ubi-1 promoter comprises a deletion of 3', 5', or both HSEs, replacement of HSEs with a trimer of a seed specific element from the promoter of pea lectin gene *psl*, or insertion of a transcription factor binding site in the HSE region. An expression construct comprising modified Ubi-1 promoter is useful for causing expression of a structural gene (agronomic genes) or open reading frame in a plant cell. The modified Ubi-1 promoter increases expression levels beyond those observed with native ubiquitin promoter. The present sequence is maize wild-type Ubi-1 promoter heat shock element.

Sequence 25 BP; 3 A; 10 C; 6 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 6; Length 25;

Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTGAGAGTCGGCT 25

Db 654 CTGGACCCCTCTGAGAGTCGGCT 678

RESULT 3  
ACC78118  
ID ACC78118 standard; DNA; 813 BP.

XX  
AC  
ACC78118;  
XX  
DT 18-AUG-2003 (first entry)

XX  
DE Maize ubiquitin-1 promoter upstream activating region (UAR).

XX  
KW Nematode; SCPI; UCP3; SUP; nematocidal; gene therapy; promoter;

XX  
OS Zea maya.

PN WO2003033651-A2.  
XX  
PD 24-APR-2003.  
XX  
PP 08-OCT-2002; 2002WO-US032059.  
XX  
PR 16-OCT-2001; 2001US-0326667P.  
XX  
(PION-) PIONEER HI-BRED INT INC.  
XX  
PT Hu X, Lu G, Ruff RL, Schuh W;  
XX  
DR WPI; 2003-393512/37.

XX  
PT New nucleic acid molecule comprising a heterologous nematode-resistance sequence operably linked to a promoter capable of driving transcriptions of the sequence, useful for promoting nematode-resistance in plants.

XX  
PS Claim 16; Page 55; 65pp; English.

XX  
CC The invention relates to a nucleic acid molecule comprising a

CC heterologous nematode-resistance sequence operably linked to a promoter capable of driving transcriptions of the heterologous nematode-resistance

CC sequence in a plant cell, where the promoters are selected from synthetic SCPI, UCP3 and SUP promoters. The nucleic acid molecules and methods are

CC useful for promoting nematode-resistance in plants by inducing expression CC of a heterologous nematode-resistance sequence, altering expression of a heterologous nucleotide sequence, or creating or enhancing disease resistance in a plant. The present sequence represents the upstream CC activating region of (UAR) of the ubiquitin-1 promoter from maize

XX  
SQ Sequence 813 BP; 241 A; 166 C; 142 G; 264 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 8; Length 813;

Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTGAGAGTCGGCT 25

Db 654 CTGGACCCCTCTGAGAGTCGGCT 678

RESULT 4  
ACC78118  
ID ACC78118 standard; DNA; 813 BP.

XX  
AC  
ACC78118;  
XX  
DT 18-AUG-2003 (first entry)

XX  
DE Maize ubiquitin-1 promoter upstream activating region (UAR).

XX  
KW Nematode; SCPI; UCP3; SUP; nematocidal; gene therapy; promoter;

XX  
OS Zea maya.

PN WO2003033651-A2.  
XX  
PD 24-APR-2003.  
XX  
PP 08-OCT-2002; 2002WO-US032059.  
XX  
PR 16-OCT-2001; 2001US-0326667P.  
XX  
(PION-) PIONEER HI-BRED INT INC.  
XX  
PT Hu X, Lu G, Ruff RL, Schuh W;  
XX  
DR WPI; 2003-393512/37.

XX  
PT New nucleic acid molecule comprising a heterologous nematode-resistance sequence operably linked to a promoter capable of driving transcriptions of the sequence, useful for promoting nematode-resistance in plants.

XX  
PS Claim 16; Page 55; 65pp; English.

XX  
CC The invention relates to a nucleic acid molecule comprising a

CC heterologous nematode-resistance sequence operably linked to a promoter capable of driving transcriptions of the heterologous nematode-resistance

CC sequence in a plant cell, where the promoters are selected from synthetic SCPI, UCP3 and SUP promoters. The nucleic acid molecules and methods are

CC useful for promoting nematode-resistance in plants by inducing expression CC of a heterologous nematode-resistance sequence, altering expression of a heterologous nucleotide sequence, or creating or enhancing disease resistance in a plant. The present sequence represents the upstream CC activating region of (UAR) of the ubiquitin-1 promoter from maize

XX  
SQ Sequence 813 BP; 241 A; 166 C; 142 G; 264 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 8; Length 813;

Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTGAGAGTCGGCT 25

Db 654 CTGGACCCCTCTGAGAGTCGGCT 678

XX	OS	Synthetic.
XX	OS	Cauliflower mosaic virus.
XX	OS	Zea mays.
XX	PN	WO-003033651-A2.
XX	PD	24-APR-2003.
XX	PP	08-OCT-2002; 2002WO-US032059.
XX	PR	16-OCT-2001; 2001US-0329667P.
XX	PA	(PION-) PIONEER HI-BRED INT INC.
XX	XX	Hu X, Lu G, Ruff RL, Schuh W;
XX	PI	WPI; 2003-393512/37.
XX	DR	
XX	PS	Claim 1; Page 54-55; 65pp; English.
XX	CC	The invention relates to a nucleic acid molecule comprising a heterologous nematode-resistance sequence operably linked to a promoter capable of driving transcriptions of the heterologous nematode-resistance sequence in a plant cell, where the promoters are selected from synthetic SCPI, UCP3 and SUP promoters. The nucleic acid molecules and methods are useful for promoting nematode-resistance in plants by inducing expression of a heterologous nematode-resistance sequence, altering expression of a heterologous nucleotide sequence, or creating or enhancing disease resistance in a plant. The present sequence represents a nematode-regulated promoter SUP1, a synthetic hybrid promoter consisting of one copy of the upstream activating region (UAR) of the 35S CaMV promoter from Cauliflower mosaic virus operably linked to the Rsyn element which CC is operably linked to one copy of the UAR of ubiquitin-1 promoter from Z. CC may, which is operably linked to the Rsyn-Syn II core promoter.
XX	SQ	Sequence 1413 BP; 419 A; 309 C; 280 G; 405 T; 0 U; 0 Other; Query Match 100.0%; Score 25; DB 8; Length 1413; Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 25; Conservative 0; Matches 25; Mismatches 0; Indels 0; Gaps 0;
XX	Db	1086 CTGGACCCCTCTCGAGACTTCGCT 1110
XX	RESULT 5	RESULT 6
ID ACC78116	AAZ11622	AAZ11622
ID ACC78116 standard; DNA; 1413 BP.	ID AAZ11622 standard; DNA; 1807 BP.	ID AAZ11622;
XX	AC	AAZ11622;
AC ACC78116;	XX	DT
XX	XX	16-NOV-1999 (first entry)
XX	DB	UCP2 promoter sequence.
XX	XX	Plant promoter; TATA motif; transcription start site; upstream element; gene expression; oxalate oxidase; plant resistance; pathogen; maize; Ubi-1 promoter; upstream activating region; ss.
XX	OS	Synthetic.
XX	OS	Zea sp.
XX	XX	Nematode; SCP1; SUP; nematocidal; gene therapy; promoter; SCP1 promoter; ubiquitin-1; Rsyn-Syn II core promoter; ss.
XX	PN	WO9943838-A1.
KW 35S promoter;	XX	PD 02-SEP-1999.

PF 23-FEB-1999; 99WO-US003863.  
 XX  
 PR 24-FEB-1998; 98US-00028819.  
 XX  
 PA (FIION-) PIONNER HI-BRED INT INC.  
 XX  
 PI Bowen BA, Bruce WB, Lu G, Sims LE, Tagliani LA;  
 XX  
 DR WPT; 1999-540601/45.

PT New synthetic promoter functional in plants to provide non-tissue specific, constitutive expression, particularly of oxalate oxidase for increased resistance to pathogens.  
 XX Disclosure; Page 53-54; 61pp; English.  
 CC  
 CC The invention provides a new synthetic plant promoter that comprises a TATA motif; a transcription start site (TSS) and a region between TATA and TSS containing at least 64 percent GC content. The synthetic core promoter, optionally containing additional upstream elements are used to increase expression, provides non-tissue specific, constitutive transcription of heterologous genes in any sort of plant, especially the gene for oxalate oxidase for increasing plant resistance to pathogens. The upstream activating elements can be used to increase transcription from any promoter. A combination of the synthetic core promoter with synthetic upstream elements can induce expression 10 times greater than that provided by the maize Ubi-1 promoter. The present sequence represents a UCP2 promoter sequence (2 copies of Ubi1 upstream activating region (UAR) operably linked with the core promoter)

SQ Sequence 1807 BP; 530 A; 375 C; 323 G; 579 T; 0 U; 0 Other;  
 CC Query Match 100.0%; Score 25; DB 2; Length 1807;  
 CC Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;  
 CC Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC Qy 1 CTGGACCCCTCTGAGAGTCGGCT 25  
 Db 654 CTGGACCCCTCTGAGAGTCGGCT 678

RESULT 7  
 ADC4542  
 ID ADC4542 Standard; DNA; 2005 BP.  
 XX  
 AC ADC84542;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE Modified maize ubiquitin promoter encoding sequence.  
 XX  
 KW expression cassette; ds.  
 XX  
 OS Zea mays.  
 XX  
 PN WO2003027257-A2.  
 XX  
 PD 26-JUN-2003.  
 XX  
 PR 16-DEC-2002; 2002WO-US040099.  
 XX  
 PR 17-DEC-2001; 2001US-0341456P.  
 PR 08-JAN-2002; 2002US-0346660P.  
 XX  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Steiner H, Dawson J, Dunder E, Meghji M;  
 XX  
 DR WPI; 2003-533014/50.

XX  
 XX  
 PR 27-SEP-2002; 2002WO-US030895.  
 XX  
 PR 27-SEP-2001; 2001US-0325607P.  
 XX  
 PD 03-APR-2003.  
 XX  
 PR  
 XX  
 PR 27-SEP-2001; 2001US-0325607P.  
 XX  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PT New isolated nucleic acid molecule encoding a polypeptide with a 3'-5' exonuclease domain, useful in molecular biology and transformation, in particular for reproducing and predictably manipulating gene expression in a plant cell.  
 XX  
 WPI; 2003-354651/33.  
 XX  
 PT New isolated nucleic acid molecule encoding a polypeptide with a 3'-5' exonuclease domain, useful in molecular biology and transformation, in particular for reproducing and predictably manipulating gene expression in a plant cell.

XX Disclosure; Page 105-106; 108pp; English.

XX  
 CC The present invention relates to an invention that results in the expression of the endogenous nucleotide sequence in plant cell being increased. In this method the plant cell comprises a second expression cassette comprising a nucleic acid molecule, where the expression of the nucleic acid molecule of interest in the plant cell is decreased as compared to expression of the nucleic acid molecule of interest in a plant cell lacking the first expression cassette. The methods and compositions of the present invention are useful in molecular biology and transformation, in particular for reproducing and predictably manipulating gene expression in a plant cell. The present sequence is a nucleotide sequence of modified maize ubiquitin promoter.

SQ Sequence 2005 BP; 474 A; 426 C; 412 G; 693 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 10; Length 2005;  
 ID ACF06306 Standard; DNA; 2100 BP.  
 XX  
 AC ACF06306;  
 XX  
 DT 07-OCT-2003 (first entry)  
 XX  
 DB 5' genome+intact vi3A insert sequence SEQ ID NO:6.  
 XX  
 KW Maize; corn; VIP1034; vi3A; pat; vegetative insecticidal protein; VIP; transgenic genotype; insect resistance; herbicide tolerance; corn event;  
 KW plant; gene; ds.  
 XX  
 OS Zea mays.  
 XX  
 PN WO2003052073-A2.

XX  
 PD 26-JUN-2003.  
 XX  
 PR 17-DEC-2001; 2001US-0341456P.  
 PR 08-JAN-2002; 2002US-0346660P.  
 XX  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX  
 PI Steiner H, Dawson J, Dunder E, Meghji M;

XX  
 DR WPI; 2003-533014/50.

XX  
 PT New transgenic corn event, VIP1034, useful for conferring a characteristic consisting of insect resistance and herbicide tolerance.  
 XX  
 PS Claim 7; Page 56-57; 73pp; English.

XX  
 CC The present invention describes an isolated nucleic acid molecule (I) comprising a sequence having at least one junction sequence of corn event (II) comprising a sequence having 20 bp, or its complement. VIP1034 is a vegetative insecticidal protein (ViP), and comprises a novel transgenic genotype that comprises a vi3A gene and a pat gene, which confers insect resistance and herbicide tolerance to the transgenic corn event. Also described: (1) an amplicon comprising (I); (2) a polynucleotide primer for detecting corn event VIP1034 in a sample comprising at least 15 contiguous nucleotides from position 1-716 of the 2100-bp sequence given in ACF06306 or its complement; (3) detecting the presence of a DNA corresponding to the VIP1034 event in the sample; (4) a kit for detecting

CC the presence of VIP1034 nucleic acids in a sample; (5) detecting corn event VIP1034 in a biological sample; (6) a seed of any corn inbred comprising the transgenic genotype of the corn event VIP1034, where the transgenic genotype comprises an intact copy and a fragmented copy of a first expression cassette and an intact copy of a second expression cassette; (7) a corn plant or its pollen or ovule comprising the transgenic genotype of the corn event VIP1034; (8) a hybrid corn seed produced by crossing the corn plant with an inbred corn plant having a different genotype; (9) a hybrid corn plant produced by growing the hybrid corn seed; (10) producing corn seed; and (11) producing hybrid corn seeds. The nucleic acid is useful for conferring a characteristic consisting of insect resistance and herbicide tolerance. The present sequence represents a corn 5' genome+intact VIP3A insert sequence, which is used in the exemplification of the present invention

XX Sequence 2100 BP; 555 A; 479 C; 397 G; 668 T; 0 U; 1 Other;

SQ Query Match 100.0%; Score 25; DB 9; Length 2100;  
Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;

DB Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTCGAGAGTCGGCT 25

Db 1401 CTGGACCCCTCTCGAGAGTCGGCT 1425

RESULT 9

ACCT8115  
ID ACC78115 standard; DNA; 2609 BP.

AC ACCT8115;  
XX DT 18-AUG-2003 (first entry)

DE Nematode-regulated synthetic promoter UCP3.

XX KW Nematode; SCP1; UCP3; SUP; nematocidal; gene therapy; promoter;  
KW ubiquitin-1; Rsyn7-Syn II core promoter; ss.  
XX OS Synthetic.  
OS Zea mays.  
XX PN WO2003033651-A2.  
XX PD 24-APR-2003.  
XX PR 16-OCT-2002; 2002WO-US032059.  
XX PA (PION-) PIONEER HI-BRED INT INC.  
XX PI Bowen BA, Bruce WB, Lu G, Sims LE, Tagliani LA;  
XX DR WPI; 1999-540601/45.  
XX PT New synthetic promoter functional in plants to provide non-tissue specific, constitutive expression, particularly of oxalate oxidase for increased resistance to pathogens.  
XX PS Disclosure; Page 54-56; 61pp; English.

XX CC The invention provides a new synthetic plant promoter that comprises a TATA motif; a transcription start site (TSS) and a region between TATA and TSS containing at least 64 percent GC content. The synthetic core promoter, optionally containing additional upstream elements are used to increase expression. Provides non-tissue specific, constitutive transcription of heterologous genes in any sort of plant, especially the gene for oxalate oxidase for increasing plant resistance to pathogens. The upstream activating elements can be used to increase transcription from any promoter. A combination of the synthetic core promoter with synthetic upstream elements can induce expression 10 times greater than that provided by the maize Ubi1 promoter. The present sequence represents a copy of Ubi1 upstream activating region (UAR) operably linked with the core promoter

XX Sequence 2620 BP; 772 A; 541 C; 465 G; 842 T; 0 U; 0 Other;  
Query Match 100.0%; Score 25; DB 2; Length 2620;  
Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTCGAGAGTCGGCT 25

CC copies (in forward orientation) of the upstream activating region (UAR) of the ubiquitin-1 promoter from *Z. mays* operably linked to the Rsyn7-Syn II core promoter

XX Sequence 2609 BP; 769 A; 542 C; 464 G; 834 T; 0 U; 0 Other;

SQ Query Match 100.0%; Score 25; DB 8; Length 2609;  
Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;

DB Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTCGAGAGTCGGCT 25

Db 654 CTGGACCCCTCTCGAGAGTCGGCT 678

RESULT 10

AAZ11623  
ID AAZ11623 standard; DNA; 2620 BP.

XX AC AAZ11623;  
XX DT 16-NOV-1999 (first entry)

DE Ubi1 upstream activating region (UAR) linked with a promoter fragment.

XX KW Plant promoter; TATA motif; transcription start site; upstream element; KW gene expression; oxalate oxidase; plant resistance; pathogen; maize;  
KW Ubi-1 promoter; upstream activating region; ss.  
XX OS Synthetic.  
OS Zea sp.

XX PN WO9943838-A1.

XX PD 02-SEP-1999.

XX PP 23-FEB-1999; 99WO-US003863.

XX PR 24-FEB-1998; 98US-00028819.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Bowen BA, Bruce WB, Lu G, Sims LE, Tagliani LA;

XX DR WPI; 1999-540601/45.

XX PT New synthetic promoter functional in plants to provide non-tissue specific, constitutive expression, particularly of oxalate oxidase for increased resistance to pathogens.

XX PS Disclosure; Page 54-56; 61pp; English.

XX CC The invention provides a new synthetic plant promoter that comprises a TATA motif; a transcription start site (TSS) and a region between TATA and TSS containing at least 64 percent GC content. The synthetic core promoter, optionally containing additional upstream elements are used to increase expression. Provides non-tissue specific, constitutive transcription of heterologous genes in any sort of plant, especially the gene for oxalate oxidase for increasing plant resistance to pathogens. The upstream activating elements can be used to increase transcription from any promoter. A combination of the synthetic core promoter with synthetic upstream elements can induce expression 10 times greater than that provided by the maize Ubi1 promoter. The present sequence represents a copy of Ubi1 upstream activating region (UAR) operably linked with the core promoter

XX Sequence 2620 BP; 772 A; 541 C; 465 G; 842 T; 0 U; 0 Other;  
Query Match 100.0%; Score 25; DB 2; Length 2620;  
Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTCGAGAGTCGGCT 25

Db 654 |||CTGGACCCCTCTCGAGAGTTCGCT 678 AC ACF06318;  
 RESULT 11 XX  
 AAH19412 standard; DNA; 2807 BP. DT 07-OCT-2003 (first entry)  
 ID XX DE Pat insert nucleotide sequence SEQ ID NO:18.  
 AAH19412; KW Maize; corn; VIP1034; vip3A; pat; vegetative insecticidal protein; VIP;  
 XX KW transgenic genotype; insect resistance; herbicide tolerance; corn event;  
 DT 30-JUL-2001 (first entry) KW plant; gene; ds.  
 XX OS Zea mays.  
 DE SUC-1 promoter element. XX  
 XX KW Promoter; SUC-1; transgenic plant; CaMV 35S promoter; maize;  
 KW polyubiquitin promoter; UBI promoter; sugarcane gene expression;  
 XX monocoyledonous plant gene expression; ds.  
 OS Synthetic. XX  
 XX  
 FH Location/Qualifiers PR WO2003052073-A2.  
 FT 5'UTR 1715..1797 PR 08-JAN-2002; 2002US-034660P.  
 FT /\*tag= a XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 FT /note= "Maize UBI 5' UTR" XX  
 FT 1798..2807 PR 16-DEC-2002; 2002WO-US040099.  
 FT /\*tag= c XX DR Steiner H, Dawson J, Dunder E, Meghji M;  
 XX PN WO200132897-A2. XX DR 26-JUN-2003.  
 XX PD 10-MAY-2001. XX PR 06-NOV-2000; 2000WO-1B001601.  
 XX PR 05-NOV-1999; 99ZA-00006954. XX PR 11-MAY-2000; 2000ZA-00002313.  
 PA (SASU-) SOUTH AFRICAN SUGAR ASSOC. CC VIP1034 comprising a sequence having at least one junction sequence of corn event  
 PA (URST-) UNIV STELLENBOSCH INST PLANT BIOTECHNOLO. CC VIP1034 comprising a sequence having 20 bp, or its complement. VIP1034 is  
 XX a vegetative insecticidal protein (VIP), and comprises a novel transgenic  
 PI Greenewald J, Botha FC; CC genotype that comprises a VIP3A gene and a pat gene, which confers insect  
 XX resistance and herbicide tolerance to the transgenic corn event. Also  
 DR WPI; 2001-316449/33. CC described: (1) an amplicon comprising (1); (2) a polynucleotide primer  
 XX A new promoter construct designated SUC-1 contains promoter elements from CC for detecting corn event VIP1034 in a sample comprising at least 15  
 PT CAMV 35S and maize UBI and is useful to confer stable high level CC contiguous nucleotides from position 1-716 of the 2100-bp sequence given  
 PT constitutive expression in transgenic plants. CC in ACF06306 or its complement; (3) detecting the presence of a DNA  
 PS Claim 1; Fig 1; 19pp; English. CC corresponding to the VIP1034 event in the sample; (4) a kit for detecting  
 PT The present sequence is tandem promoter construct, SUC-1. This promoter CC the presence of VIP1034 nucleic acids in a sample; (5) detecting corn  
 CC confers stable, high level, constitutive expression in transgenic plants. CC event VIP1034 in a biological sample; (6) a seed of any corn inbred  
 CC The promoter contains both the cauliflower mosaic virus (CaMV) 35S CC comprising the transgenic genotype of the corn event VIP1034, where the  
 CC promoter and the maize polyubiquitin (UBI) promoter. The promoter is CC transgenic genotype comprises an intact copy and a fragmented copy of a  
 CC particularly useful for gene expression in monocotyledonous plants, such CC first expression cassette and an intact copy of a second expression  
 CC as sugarcane. The combination of the two promoters gives higher CC cassette; (7) a corn plant or its pollen or ovule comprising the  
 CC expression levels in monocotyledonous plants than the CaMV 35S or UBI CC transgenic genotype of the corn event VIP1034; (8) a hybrid corn seed  
 CC promoter when used alone. CC produced by crossing the corn plant with an inbred corn plant having a  
 XX Sequence 2807 BP; 751 A; 607 C; 582 G; 867 T; 0 U; 0 Other; CC different genotype; (9) a hybrid corn plant produced by growing the  
 XX Best Local Similarity 100.0%; Score 25; DB 4; Length 2807; CC hybrid corn seed; (10) producing corn seed; and (11) producing hybrid  
 XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0; CC corn seeds. The nucleic acid is useful for conferring a characteristic  
 QY 1 CTGGACCCCTCTCGAGAGTTCGCT 25 CC sequence consisting of insect resistance and herbicide tolerance. The present  
 Db 1501 CTGGACCCCTCTCGAGAGTTCGCT 1525 CC sequence represents a corn pat insert nucleotide sequence, which is used  
 XX Sequence 2848 BP; 937 A; 632 C; 580 G; 699 T; 0 U; 0 Other;  
 XX Query Match 100.0%; Score 25; DB 9; Length 2848;  
 XX Best Local Similarity 100.0%; Score 25; DB 9; Length 2848;  
 XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTGGACCCCTCTCGAGAGTTCGCT 25  
 Db 2159 CTGGACCCCTCTCGAGAGTTCGCT 2135

RESULT 12 ACF06318/C  
 ID ACF06318 standard; DNA; 2848 BP.  
 XX

RESULT 13 AAD11579  
 ID AAD11579 standard; DNA; 3069 BP.  
 XX

Query Match		Score	DB	Length	Matches
		100.0%	4	3069	25
XX	AADI1579;				
XX	24-SEP-2001 (first entry)				
DE	pND3018 repression construct.				
KW	zinc finger protein; ZFP; gene expression; metabolic pathway regulator; modulation; plant technology; agriculture; pND3018 repression construct; ds.				
XX	Unidentified.				
OS					
FH	Location/Qualifiers				
FT	Key promoter	44. .2026			
FT	/tag= a				
FT	/label= 2mUbi promoter				
FT	misc_feature	2066. .2173			
FT	/tag= b				
FT	/note= "Sins interaction domain (SID) repression domain"				
FT	protein_bind	2207. .2235			
FT	/tag= d				
FT	/bound_moiety= "ZFP"				
FT	big_peptide	2620. .2641			
FT	/tag= C				
FT	/note= "nuclear localisation signal"				
FT	misc_feature	2762. .2791			
FT	/tag= e				
FT	/note= "HA epitope tag"				
FT	terminator	2820. .3112			
FT	/tag= f				
FT	/label= Nos_terminator				
PN	WO200152620-A2.				
XX					
PD	26-JUL-2001.				
XX					
PR	19-JAN-2001; 2001WO-US001817.				
PR	21-JUL-2000; 2000US-00620897.				
XX					
PA	(Scripps Res Inst.) SCRIPPS RES INST.				
PA	(SYGN ) SYNGENTA AGRIC DISCOVERY INC.				
XX					
PI	Barbas CF, Stege JT, Guan X, Dalmia B;				
XX					
DR	WPI: 2001-465325/50.				
XX					
PT	New zinc finger proteins, useful for modulating or regulating gene expression and metabolic pathways in plants, e.g. for treating in the plant cells a disorder that is associated with abnormal expression of the target gene.				
XX					
PS	Claim 121; Page 134-136; 156pp; English.				
XX					
CC	The patient discloses methods and compositions to modulate the expression of a target gene in plant cells. The method involves providing plant cells with a zinc finger protein (ZFP) which is capable of specifically binding to a target nucleotide sequence or its complementary strand within a target gene and allowing the ZFP binding to the target nucleotide sequence, where the expression of the target gene in the plant cells is modulated. The ZFP and fusions of the ZFP proteins are useful for modulating or regulating gene expression and metabolic pathways in plants. The ZFP, fusion proteins and methods are useful in plant and agricultural technology. The method is useful particularly for treating a disorder in the plant cells, where the disorder is associated with abnormal expression of the target gene. The present DNA sequence is pND3018 repression construct. This sequence is used to construct maize ubiquitin promoter zmbf effector construct which is further used to determine the function of ZFP-effector fusion protein on luciferase reporter gene in maize cells				
CC	Sequence 3069 BP; 747 A; 713 C; 669 G; 940 T; 0 U; 0 Other;				
XX					
XX	RESULT 14				
ID	AADI1578				
XX					
AC	AADI1578;				
XX					
DT	24-SEP-2001 (first entry)				
XX					
DR	pND3008 activation construct.				
XX					
KW	zinc finger protein; ZFP; gene expression; metabolic pathway regulator; modulation; plant technology; agriculture; pND3008 activation construct; ds.				
XX	Unidentified.				
OS					
FH	Location/Qualifiers				
FT	Key promoter	44. .2026			
FT	/tag= a				
FT	/label= 2mUbi_promoter				
FT	misc_feature	2066. .2173			
FT	/tag= b				
FT	/note= "Sins interaction domain (SID) repression domain"				
FT	protein_bind	2207. .2235			
FT	/tag= d				
FT	/bound_moiety= "ZFP"				
FT	big_peptide	2620. .2641			
FT	/tag= C				
FT	/note= "nuclear localisation signal"				
FT	misc_feature	2762. .2791			
FT	/tag= e				
FT	/note= "HA epitope tag"				
FT	terminator	2820. .3112			
FT	/tag= f				
FT	/label= Nos_terminator				
PN	WO200152620-A2.				
XX					
PD	26-JUL-2001.				
XX					
PR	19-JAN-2001; 2001WO-US001817.				
PR	21-JAN-2000; 2000US-00620897.				
PR	21-JUL-2000; 2000US-00620897.				
XX					
PA	(Scripps Res Inst.) SCRIPPS RES INST.				
PA	(SYGN ) SYNGENTA AGRIC DISCOVERY INC.				
XX					
PI	Barbas CF, Stege JT, Guan X, Dalmia B;				
XX					
DR	WPI: 2001-465325/50.				
XX					
PT	New zinc finger proteins, useful for modulating or regulating gene expression and metabolic pathways in plants, e.g. for treating in the plant cells a disorder that is associated with abnormal expression of the target gene.				
CC	Sequence 3069 BP; 747 A; 713 C; 669 G; 940 T; 0 U; 0 Other;				
CC	Example 1B; Page 132-134; 156pp; English.				
CC	The patient discloses methods and compositions to modulate the expression of a target gene in plant cells. The method involves providing plant cells with a zinc finger protein (ZFP) which is capable of specifically binding to a target nucleotide sequence or its complementary strand				
CC	binding to a target nucleotide sequence or its complementary strand				

CC within a target gene and allowing the ZFP binding to the target nucleotide sequence, where the expression of the target gene in the plant cells is modulated. The ZFP and fusions of the ZFP proteins are useful for modulating or regulating gene expression and metabolic Pathways in plants. The ZFP, fusion proteins and methods are useful in plant and agricultural technology. The method is useful particularly for treating a disorder in the plant cells, where the disorder is associated with abnormal expression of the target gene. The present DNA sequence is PND308 activation construct. This sequence is used to construct maize ubiquitin promoter ZmUbi effector construct which is further used to determine the function of ZFP-effector fusion protein on luciferase reporter gene in maize cells

SQ

Sequence 3121 BP; 735 A; 713 C; 684 G; 969 T; 0 U; 0 Other;

Query Match

Best Local Similarity 100.0%; Score 25; DB 4; Length 3121; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 CTGGACCCCTCTGAGAGTTGGCT 25

Db

677 CTGGACCCCTCTGAGAGTTGGCT 701

Search completed: February 28, 2005, 23:51:37

Job time : 223.182 Secs

CC region (UAR) operably linked with the core promoter)

XX Sequence 3433 BP; 1014 A; 707 C; 607 G; 1105 T; 0 U; 0 Other;

Query Match Best Local Similarity 100.0%; Score 25; DB 2; Length 3433; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTGAGAGTTGGCT 25

Db 654 CTGGACCCCTCTGAGAGTTGGCT 678

RESULT 15

AAZ11624 Standard; DNA; 3433 BP.

ID

AAZ11624;

XX

AC 16 NOV-1999 (first entry)

DT

DE UCP4 promoter sequence.

XX

KW Plant promoter; TATA motif; transcription start site; upstream element;

KW

Ubi-1 promoter; upstream activating region; 5S. gene expression; oxalate oxidase; plant resistance; pathogen; maize; OS Synthetic.

OS

Zea sp.

XX

W09943838-A1.

PN

02-SEP-1999.

PD

XX 23-FEB-1999; 99WO-US003863.

PF

XX 24-FEB-1998; 98US-00028819.

PR

PA (PION-) PIONER HI-BRED INT INC.

XX

PT Bowen BA, Bruce WB, Lu G, Sims LE, Tegliani LA;

PT

XX DR WPI; 1999-540601/45.

PS Disclosure; Page 56-58; 61PP; English.

XX

PT New synthetic promoter functional in plants to provide non-tissue specific, constitutive expression, particularly of oxalate oxidase for increased resistance to pathogens.

XX Disclosure; Page 56-58; 61PP; English.

CC The invention provides a new synthetic plant promoter that comprises a TATA motif; a transcription start site (TSS) and a region between TATA and TSS containing at least 64 percent GC content. The synthetic core promoter, optionally containing additional upstream elements are used to increase expression. Provides non-tissue specific, constitutive transcription of heterologous genes in any sort of plant, especially the gene for oxalate oxidase for increasing plant resistance to pathogens.

CC The upstream activating elements can be used to increase transcription from any promoter. A combination of the synthetic core promoter with synthetic upstream elements can induce expression 10 times greater than CC that provided by the maize Ubi-1 promoter. The present sequence represents a UCP4 promoter sequence (4 copies of Ubi1 upstream activating

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:35:01 ; Search time 1734.42 Seconds  
 (without alignments)  
 548.661 Million cell updates/sec

Title: US-10-086-062-1

Perfect score: 25

Sequence: 1 ctggacccttcgagatccgt 25

Scoring table: IDENTITY\_NUC

Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
 Listing first 45 summaries

Database : EST\*

1: gb\_eat1:\*

2: gb\_eat2:\*

3: gb\_htc:\*

4: gb\_eat3:\*

5: gb\_eat4:\*

6: gb\_eat5:\*

7: gb\_eat6:\*

8: gb\_gb81:\*

9: gb\_gb82:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

## ALIGNMENTS

25	25	100.0	859	7	CO47274
26	25	100.0	876	7	CO47281
27	25	100.0	879	7	CO47271
28	25	100.0	940	9	CC670569
c	29	100.0	965	9	CG342774
	30	23	92.0	7	CO472764
	31	23	92.0	7	CO472754
	32	21	84.0	7	CK438248
	33	20.2	80.8	100	BH219170
	34	20.2	80.8	415	BM61080
	35	20.2	80.8	553	BN736979
	36	20.2	80.8	636	BI679714
c	37	20.2	80.8	728	CQ212795
	38	20.2	80.8	896	CG212788
	39	19.8	79.2	300	C55594
	40	19.2	75.8	669	CL173701
	41	19.2	76.8	897	BU838185
	42	19.2	76.8	908	BB797311
	43	19.2	75.8	1283	AZ130054
c	44	18.8	75.2	323	B102971
c	45	18.8	75.2	472	BE142500

RESULT 1

BH638797

LOCUS

BH638797

DEFINITION

1008024911-2BL\_x1

1008

-

229 bp

DNA

linear

GSS

14-FEB-2002

survey sequence.

ACCESSION

BH638797

VERSION

GI:18663263

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Zea.

1 (bases 1 to 229)

Class: transposon-tagged.

Location/Qualifiers

1. .229

/FEATURES

source

/organism="Zea mays"

/mol\_type="genomic DNA"

/cultivar="mixed background

/db\_xref="taxon:4577"

/db\_xref="

and BgIII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

## ORIGIN

QY	Db	RESULT 2	RESULT 3
CG805015	CG805015	Query Match Best Local Similarity 100.0%; Score 25; DB 9; Length 307; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	CG805036 Query Match Best Local Similarity 100.0%; Score 25; DB 9; Length 338; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	DEFINITION CG805015	DEFINITION CG805036
QY	Db	LOCUS CG805015	LOCUS CG805036
QY	Db	ACCESSION CG805015	ACCESSION CG805036
QY	Db	VERSION CG805015.1	VERSION CG805036.1
QY	Db	KEYWORDS GSS.	KEYWORDS GSS.
QY	Db	SOURCE Zea mays	SOURCE Zea mays
QY	Db	ORGANISM Zea mays	ORGANISM Zea mays
QY	Db	Bakteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD	Bakteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

RESULT 2  
Query Match Best Local Similarity 100.0%; Score 25; DB 9; Length 229; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
LOCUS 1118056D09\_2EL\_X1 1118 - Rescuemu Grid S Zea mays genomic, genomic survey sequence.  
DEFINITION 1118056F03\_2EL\_X1 1118 - Rescuemu Grid S Zea mays genomic, genomic survey sequence.  
ACCESSION CG805015  
VERSION CG805015.1  
KEYWORDS GSS.  
SOURCE Zea mays

RESULT 2  
Query Match Best Local Similarity 100.0%; Score 25; DB 9; Length 229; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
LOCUS 1118056D09\_2EL\_X1 1118 - Rescuemu Grid S Zea mays genomic, genomic survey sequence.  
DEFINITION 1118056F03\_2EL\_X1 1118 - Rescuemu Grid S Zea mays genomic, genomic survey sequence.  
ACCESSION CG805015  
VERSION CG805015.1  
KEYWORDS GSS.  
SOURCE Zea mays

COMMENT Contact: Walbot V Department of Biological Sciences

RESULTS

FEATURES	source	FEATURES	source
1.	.338	1.	.338
/organism="Zea mays"	/organism="Zea mays"	/mol_type="genomic DNA"	/mol_type="genomic DNA"
/mol_type="genomic DNA"	/mol_type="genomic DNA"	/cultivar="mixed background W23/A188/B73"	/cultivar="mixed background W23/A188/B73"
/cultivar="mixed background W23/A188/B73"	/cultivar="mixed background W23/A188/B73"	/db_xref="taxon:4577"	/db_xref="taxon:4577"
/db_xref="taxon:4577"	/db_xref="taxon:4577"	/dev_stage="adult"	/dev_stage="adult"
/tissue="leaf"	/tissue="leaf"	/lab_host="DH10B"	/lab_host="DH10B"
/clone_ljb="1118 - Rescuemu Grid S"	/clone_ljb="1118 - Rescuemu Grid S"	/clone_ljb="1118 - Rescuemu Grid S"	/clone_ljb="1118 - Rescuemu Grid S"
/note="Organ: leaf; Vector: Rescuemu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BgIII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'Rescuemu'. Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BgIII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."	/note="Organ: leaf; Vector: Rescuemu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BgIII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'Rescuemu'. Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BgIII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."	Query Match Best Local Similarity 100.0%; Score 25; DB 9; Length 338; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Query Match Best Local Similarity 100.0%; Score 25; DB 9; Length 338; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN	ORIGIN	ORIGIN	ORIGIN
RESULT 4	RESULT 4	RESULT 4	RESULT 4
BH38999	BH38999	BH38999	BH38999
LOCUS	LOCUS	LOCUS	LOCUS
DEFINITION	DEFINITION	DEFINITION	DEFINITION
100826F08_1EL_X1 1008 - Rescuemu Grid I Zea mays genomic, genomic survey sequence.	100826F08_1EL_X1 1008 - Rescuemu Grid I Zea mays genomic, genomic survey sequence.	100826F08_1EL_X1 1008 - Rescuemu Grid I Zea mays genomic, genomic survey sequence.	100826F08_1EL_X1 1008 - Rescuemu Grid I Zea mays genomic, genomic survey sequence.
ACCESSION	ACCESSION	ACCESSION	ACCESSION
BH638999	BH638999	BH638999	BH638999
VERSION	VERSION	VERSION	VERSION
BH638999.1	BH638999.1	BH638999.1	BH638999.1
KEYWORDS	KEYWORDS	KEYWORDS	KEYWORDS
GSS.	GSS.	GSS.	GSS.
SOURCE	SOURCE	SOURCE	SOURCE
Zea mays	Zea mays	Zea mays	Zea mays

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD  
1 bases 1 to 534)  
clade; Panicoideae; Andropogoneae; Zea.  
Walbot, V.  
Maize genomic sequences found using engineered Rescuemu transposon  
Unpublished (2001)  
Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Very probable ligation site of ends cut by single endonuclease.  
Reverse complemented post-ligation sequence from source sequence.  
Plate: 100826 row: 10  
Class: transposon-tagged  
Location/Qualifiers  
FEATURES  
Source  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/B73"  
/db\_xref="taxon:4577"  
/tissue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="119 - Rescuemu Grid AA"  
/note="Organ: leaf; Vector: Rescuemu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site www.zmdb.iastate.edu and follow the links for 'Rescuemu'. Grid AA was grown at UC San Diego in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."  
ORIGIN  
Query Match 100.0%; Score 25; DB 8; Length 534;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGGACCCCTCTCGAGAGTTCGCT 25  
Db 246 CTGGACCCTCTCGAGAGTTCGCT 270  
RESULT 5  
CG80054  
LOCUS CG80054  
DEFINITION 11:805601..28L\_x1 1118 - Rescuemu Grid S ze a m a y s g e n o m i c \_ g e n o m i c  
ACCESSION CG80054  
VERSION 1  
KEYWORDS GSS,  
SOURCE Zea mays  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD  
REFERENCE  
TITLE Maize genomic sequences found using engineered Rescuemu transposon  
AUTHORS Walbot, V  
JOURNAL Unpublished (2001)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Possible ligation site of ends cut by 2 different endonucleases.  
Reverse complemented post-ligation sequence from source sequence.  
Plate: 1119014 row: 29  
Class: transposon-tagged  
Location/Qualifiers  
FEATURES  
Source  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/B73"  
/db\_xref="taxon:4577"  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Very probable ligation site of ends cut by single endonuclease.  
Reverse complemented post-ligation sequence from source sequence.  
Plate: 100826 row: 10  
Class: transposon-tagged  
Location/Qualifiers  
FEATURES  
Source  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/B73/  
/db\_xref="taxon:4577"  
/tissue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="119 - Rescuemu Grid AA"  
/note="Organ: leaf; Vector: Rescuemu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site www.zmdb.iastate.edu and follow the links for 'Rescuemu'. Grid AA was grown at UC San Diego in 2001. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."  
ORIGIN  
Query Match 100.0%; Score 25; DB 9; Length 617;  
Best Local Similarity 100.0%; Pred. No. 0.68;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGGACCCCTCTCGAGAGTTCGCT 25  
Db 356 CTGGACCCTCTCGAGAGTTCGCT 380  
RESULT 6  
CG80054  
LOCUS CG80054  
DEFINITION 11:805601..28L\_x1 1118 - Rescuemu Grid S ze a m a y s g e n o m i c \_ g e n o m i c  
ACCESSION CG80054  
VERSION 1  
KEYWORDS GSS,  
SOURCE Zea mays  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD  
REFERENCE  
TITLE Maize genomic sequences found using engineered Rescuemu transposon  
AUTHORS Walbot, V  
JOURNAL Unpublished (2001)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Possible ligation site of ends cut by 2 different endonucleases.  
Reverse complemented post-ligation sequence from source sequence.  
Plate: 1119056 row: 40  
Class: transposon-tagged  
Location/Qualifiers  
FEATURES  
Source  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/B73"  
/db\_xref="taxon:4577"

/tissue\_type="leaf"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="1118 - Rescuemu Grid S"  
 /note="Organ: leaf; Vector: Rescuemu (engineered from Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA units. For more information on Rescuemu, go to the website, www.zmbs-labstate.edu, and follow the links for Rescuemu." Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BgII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

## ORIGIN

Query Match 100.0%; Score 25; DB 9; Length 620;  
 Best Local Similarity 100.0%; Pred. No. 0.68; Mismatches 0; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGCCCTCTGAGAGTTCGGCT 25  
 Db 358 CTGGACCCCTCTGAGAGTTCGGCT 382

## RESULT 7

CO472745 CO472745 716 bp mRNA linear EST 09-JUL-2004  
 DEFINITION Picea glauca CDNA clone Q0081.BR\_D19 5', mRNA sequence.  
 ACCESSION Q0081.BR\_1\_D19 G0008: Cambium, phloem and bark of girdled saplings  
 VERSION CO472745  
 KEYWORDS EST  
 ORGANISM Picca glauca (white spruce)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.  
 1 (bases 1 to 76)  
 Morency, M.-J., Cooke, J., Pavly, N., Parsons, L., Paule, C., Seguin, A., Ratzel, E., Butterfield, Y., Barber, S., Yang, G., Scott, J., Suddiqui, A., Holt, R., Marr, M., and Mackay, J.

TITLE Arborea EST sequencing in Picea glauca (white spruce)  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: John Mackay  
 Pavillon Charles-Bugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
 Universite Laval  
 Centre de Recherche en Biologie Forestiere  
 Pavillon Charles-Bugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
 Fax: 418 656 7493  
 Email: j.mackay@rbsv.ulaval.ca  
 Center for Computational Genomics and Bioinformatics (CCGB), University of Minnesota, MN id Identifier: MN519339 Clone ID: G0081.BR\_D19 Clones available through: John Mackay, Ph. D. Professeur adjoint - Assistant professor EMAIL: j.mackay@rbsv.ulaval.ca Centre de Recherche en Biologie Forestiere (Forest Biology Research Center) Universite Laval Quebec, Quebec Plate: 1.BR row: 19 column: D  
 seq primer: M13 Reverse Primer

FEATURES source  
 1. .716  
 /organism="Picca glauca"  
 /mol\_type="mRNA"  
 /strain="Pg-653"  
 /db\_xref="taxon:3330"  
 /clone="G0081.BR\_D19"  
 /sex="Hermaphrodite"  
 /tissue\_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide /tissue\_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"  
 /dev\_stage="Vascular cambium, secondary phloem, and bark treatment"  
 /lab\_host="E. coli DH10B cells"  
 /clone\_lib="G0008: Cambium, phloem and bark of girdled saplings"

treatment"  
 /lab\_host="E. coli DH10B cells"  
 /clone\_lib="G0008: Cambium, phloem and bark of girdled saplings"  
 /note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBluescript II SK (+) XR; Site 1: Eco-RI; Site 2: Xba-I; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitrogen) for propagation"

## ORIGIN

Query Match 100.0%; Score 25; DB 7; length 716;  
 Best Local Similarity 100.0%; Pred. No. 0.68; Mismatches 0; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGCCCTCTGAGAGTTCGGCT 25  
 Db 666 CTGGACCCCTCTGAGAGTTCGGCT 690

## RESULT 8

CO472714 CO472714 765 bp mRNA linear EST 09-JUL-2004  
 DEFINITION Picea glauca CDNA clone Q0081.BR\_C16 5', mRNA sequence.  
 ACCESSION CO472714  
 VERSION CO472714.1 GI:50141206  
 KEYWORDS EST  
 SOURCE Picca glauca (white spruce)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.  
 1 (bases 1 to 765)  
 Morency, M.-J., Cooke, J., Pavly, N., Parsons, L., Paule, C., Seguin, A., Ratzel, E., Butterfield, Y., Barber, S., Yang, G., Scott, J., Suddiqui, A., Holt, R., Marr, M., and Mackay, J.

TITLE Arborea EST sequencing in Picea glauca (white spruce)  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: John Mackay  
 Centre de Recherche en Biologie Forestiere  
 Universite Laval  
 Pavillon Charles-Bugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
 Fax: 418 656 7493  
 Email: j.mackay@rbsv.ulaval.ca  
 Center for Computational Genomics and Bioinformatics (CCGB), University of Minnesota, MN id Identifier: MN5197864 Clone ID: G0081.BR\_C16 Clones available through: John Mackay, Ph. D. Professeur adjoint - Assistant professor EMAIL: j.mackay@rbsv.ulaval.ca Centre de Recherche en Biologie Forestiere (Forest Biology Research Center) Universite Laval Quebec, Quebec Plate: 1.BR row: 16 column: C  
 seq primer: M13 Reverse Primer

FEATURES source  
 1. .765  
 /organism="Picca glauca"  
 /mol\_type="mRNA"  
 /strain="Pg-653"  
 /db\_xref="taxon:3330"  
 /clone="G0081.BR\_C16"  
 /sex="Hermaphrodite"  
 /tissue\_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"  
 /dev\_stage="Vascular cambium, secondary phloem, and bark treatment"  
 /lab\_host="E. coli DH10B cells"  
 /clone\_lib="G0008: Cambium, phloem and bark of girdled saplings"

/note="Organ: Main stem of 4 year old saplings,  
approximately 60 cm tall; Vector: pBluescript II SK (+)  
XR; Site\_1: Eco-RI; Site\_2: Xba-I; Tissues were pooled  
from above and below the girdle. cDNA was prepared from 5  
mg of poly A+ selected RNA and was directionally ligated  
into the pBluescript II SK (+) XR vector (Stratagene),  
transformed by electroporation into DH10B cells (in  
vitrogen) for propagation"

## ORIGIN

Query Match 100.0%; Score 25; DB 7; length 765;  
Best Local Similarity 100.0%; Pred. No. 0.68; Mi. Mismatches 0;  
Matches 25; Conservative 0; Mi. Mismatches 0; Indels 0; Gaps 0;

## ORIGIN

Query Match 100.0%; Score 25; DB 7; length 769;  
Best Local Similarity 100.0%; Pred. No. 0.68; Mi. Mismatches 0;  
Matches 25; Conservative 0; Mi. Mismatches 0; Indels 0; Gaps 0;

## ORIGIN

Query Match 100.0%; Score 25; DB 7; length 769;  
Best Local Similarity 100.0%; Pred. No. 0.68; Mi. Mismatches 0;  
Matches 25; Conservative 0; Mi. Mismatches 0; Indels 0; Gaps 0;

RESULT 9  
CO472736 CO472736 769 bp mRNA linear EST 09-JUL-2004  
DEFINITION G00081.BR\_1.G06 G0008: Cambium, phloem and bark of girdled saplings  
LOCUS CO472736  
ACCESSION CO472736  
VERSION CO72736.1  
KEYWORDS EST  
SOURCE Picea glauca (white spruce)  
ORGANISM Pinaceae  
REFERENCE 1 (bases 1 to 769)  
AUTHORS Morency,M.-J., Cooke,J., Pavly,N., Parsons,L., Paule,C., Seguin,A.,  
Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J.,  
Siddiqui,A., Holt,R., Marra,M. and Mackay,J.  
TITLE Arboraea EST sequencing in Picea glauca (white spruce)  
JOURNAL Unpublished (2004)  
COMMENT Contact: John Mackay  
Centre de Recherche en Biologie Forestiere  
Universite Laval  
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
Email: jmacky@rbs.ulaval.ca  
Center for Computational Genomics and Bioinformatics (CCGB),  
University of Minnesota, MN id Identifier: MN5198051 Clone ID:  
G00081.BR\_G06 Clones available through: John Mackay, Ph. D.  
Professeur adjoint -Assistant professor EMAIL:  
jmackay@rbs.ulaval.ca Centre de Recherche en Biologie Forestiere  
(Forest Biology Research Center) Universite Laval Quebec, Quebec  
CANADA G1K 7P4  
Plate: 1.BR row: 06 column: C  
Seq primer: M13 Reverse Primer.  
FEATURES Source  
Source  
Location/Qualifiers  
1. .769  
/organism="Picea glauca"  
/mol\_type="mRNA"  
/strain="pg-653"  
/db\_xref="Taxon:3330"  
/clone="G00081.BR\_D17"  
/sex="Hermaphrodite"  
/tissue\_type="Vascular cambium, secondary phloem and bark  
of trees girdled by removing a ring of bark ca. 1 cm wide  
from the midpoint of the main stem"  
/dev\_stage="Vascular cambium, secondary phloem, and bark  
from trees harvested 1 and 7 days after girdling  
treatment"  
/lab\_host="E. coli DH10B cells"  
/clone\_lib="G0008: Cambium, phloem and bark of girdled  
saplings"  
/note="Organ: Main stem of 4 year old saplings,  
approximately 60 cm tall; Vector: pBluescript II SK (+)  
XR; Site\_1: Eco-RI; Site\_2: Xba-I; Tissues were pooled  
from above and below the girdle. cDNA was prepared from 5  
mg of poly A+ selected RNA and was directionally ligated  
into the pBluescript II SK (+) XR vector (Stratagene),  
transformed by electroporation into DH10B cells (in  
vitrogen) for propagation"

mg of poly A+ selected RNA and was directionally ligated  
into the pBluescript II SK (+) XR vector (Stratagene),  
transformed by electroporation into DH10B cells (in  
vitrogen) for propagation"

RESULT 10  
CO472678 CO472678 774 bp mRNA linear EST 09-JUL-2004  
DEFINITION G00081.BR\_1.D17 G0008: Cambium, phloem and bark of girdled saplings  
LOCUS CO472678  
ACCESSION CO472678  
VERSION CO472678.1  
KEYWORDS EST  
SOURCE Picea glauca (white spruce)  
ORGANISM Pinaceae  
REFERENCE 1 (bases 1 to 774)  
AUTHORS Morency,M.-J., Cooke,J., Pavly,N., Parsons,L., Paule,C., Seguin,A.,  
Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J.,  
Siddiqui,A., Holt,R., Marra,M. and Mackay,J.  
TITLE Arboraea EST sequencing in Picea glauca (white spruce)  
JOURNAL Unpublished (2004)  
COMMENT Contact: John Mackay  
Centre de Recherche en Biologie Forestiere  
Universite Laval  
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
Fax: 418 656 7493  
Email: jmacky@rbs.ulaval.ca  
Center for Computational Genomics and Bioinformatics (CCGB),  
University of Minnesota, MN id Identifier: MN5197337 Clone ID:  
G00081.BR\_D17 Clones available through: John Mackay, Ph. D.  
Professeur adjoint -Assistant professor EMAIL:  
jmackay@rbs.ulaval.ca Centre de Recherche en Biologie Forestiere  
(Forest Biology Research Center) Universite Laval Quebec, Quebec  
CANADA G1K 7P4  
Plate: 1.BR row: 17 column: D  
Seq primer: M13 Reverse Primer.  
FEATURES Source  
Source  
Location/Qualifiers  
1. .774  
/organism="Picea glauca"  
/mol\_type="mRNA"  
/strain="pg-653"  
/db\_xref="Taxon:3330"  
/clone="G00081.BR\_D17"  
/sex="Hermaphrodite"  
/tissue\_type="Vascular cambium, secondary phloem and bark  
of trees girdled by removing a ring of bark ca. 1 cm wide  
from the midpoint of the main stem"  
/dev\_stage="Vascular cambium, secondary phloem, and bark  
from trees harvested 1 and 7 days after girdling  
treatment"  
/lab\_host="E. coli DH10B cells"  
/clone\_lib="G0008: Cambium, phloem and bark of girdled  
saplings"  
/note="Organ: Main stem of 4 year old saplings,  
approximately 60 cm tall; Vector: pBluescript II SK (+)  
XR; Site\_1: Eco-RI; Site\_2: Xba-I; Tissues were pooled  
from above and below the girdle. cDNA was prepared from 5  
mg of poly A+ selected RNA and was directionally ligated  
into the pBluescript II SK (+) XR vector (Stratagene),  
transformed by electroporation into DH10B cells (in  
vitrogen) for propagation"

## ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 774;  
 Best Local Similarity 100.0%; Pred. No. 0.68; Mismatches 0; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGAGACCCCTCTGAGAGTTGGCT 25  
 Db 687 CTGAGACCCCTCTGAGAGTTGGCT 711

## RESULT 11

CO472728 CO472728 805 bp mRNA linear EST 09-JUL-2004  
 LOCUS GQ0081.BR\_1\_P04 GQ008: Cambium, phloem and bark of girdled saplings  
 DEFINITION Picea glauca cDNA clone GQ0081.BR\_P04 5', mRNA sequence.

ACCESSION CO472728  
 VERSION CO472728.1  
 KEYWORDS EST.  
 SOURCE Picea glauca (white spruce)  
 ORGANISM Picea glauca (white spruce)  
 REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Sporophytophyta; Coniferopsida; Coniferales; Pinaceae; Picea.  
 AUTHORS 1 (bases 1 to 805)  
 Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A.,  
 Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J.,  
 Siddiqui,A., Holt,R., Marr,M. and Mackay,J.Y.  
 TITLE Arborea EST sequencing in Picea glauca (white spruce)  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: John Mackay

Centre de Recherche en Biologie Forestiere

Universite Laval

Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
 Fax: 418 656 7493  
 Email: jmacKay@svs.ulaval.ca

Centre for Computational Genomics and Bioinformatics (CCGB),  
 University of Minnesota, MN id Identifier: MNS198005 Clone ID:  
 GQ0081.BR\_P04 Clones available through: John Mackay, Ph. D.  
 Professeur adjoint -assistant professor EMAIL:  
 jmacKay@svs.ulaval.ca Centre de Recherche en Biologie Forestiere  
 (Forest Biology Research Center) Universite Laval Quebec, Quebec  
 CANADA G1K 7P4  
 Plate: 1.BR row: 04 column: P  
 Seq primer: M13 Reverse Primer.

## FEATURES

## source

FEATURES source  
 1. .805 Location/Qualifiers  
 /organism="Picea glauca"  
 /mol\_type="mRNA"  
 /strain="PG-653"  
 /db\_xref="txon:3330"  
 /clone="GQ0081.BR\_P04"  
 /sex="hermaphrodite"  
 /tissue\_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"  
 /dev\_stage="Vascular cambium, secondary phloem and bark of trees harvested 1 and 7 days after girdling  
 /lab\_host="E. coli DH10B cells"  
 /clone\_lib="GQ008: Cambium, phloem and bark of girdled saplings"  
 /note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBlueScript II SK (+) XR; Site\_1: Eco-RI; Site\_2: Xba-I; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBlueScript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitrogen) for propagation"

## ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 805;  
 Best Local Similarity 100.0%; Pred. No. 0.69; Mismatches 0; Indels 0; Gaps 0;

## ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 817;  
 Best Local Similarity 100.0%; Pred. No. 0.69; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGAGACCCCTCTGAGAGTTGGCT 25  
 Db 695 CTGAGACCCCTCTGAGAGTTGGCT 719

## RESULT 12

CO472729 CO472729 817 bp mRNA linear EST 09-JUL-2004  
 LOCUS GQ0081.BR\_1\_E24 GQ008: Cambium, phloem and bark of girdled saplings  
 DEFINITION Picea glauca cDNA clone GQ0081.BR\_E24 5', mRNA sequence.

ACCESSION CO472729  
 VERSION CO472729.1  
 KEYWORDS EST.  
 SOURCE Picea glauca (white spruce)  
 ORGANISM Picea glauca (white spruce)  
 REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophytophyta; Coniferopsida; Coniferales; Pinaceae; Picea.  
 AUTHORS 1 (bases 1 to 817)  
 Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A.,  
 Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J.,  
 Siddiqui,A., Holt,R., Marr,M. and Mackay,J.Y.  
 TITLE Arborea EST sequencing in Picea glauca (white spruce)  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: John Mackay  
 Centre de Recherche en Biologie Forestiere  
 Universite Laval  
 Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
 Fax: 418 656 7493  
 Email: jmacKay@svs.ulaval.ca  
 Center for Computational Genomics and Bioinformatics (CCGB),  
 University of Minnesota, MN id Identifier: MNS198011 Clone ID:  
 GQ0081.BR\_E24 Clones available through: John Mackay, Ph. D.  
 Professeur adjoint -assistant professor EMAIL:  
 jmacKay@svs.ulaval.ca Centre de Recherche en Biologie Forestiere  
 (Forest Biology Research Center) Universite Laval Quebec, Quebec  
 CANADA G1K 7P4  
 Plate: 1.BR row: 24 column: E  
 Seq primer: M13 Reverse Primer.

## FEATURES

## source

FEATURES source  
 1. .817 Location/Qualifiers  
 /organism="Picea glauca"  
 /mol\_type="mRNA"  
 /strain="PG-653"  
 /db\_xref="txon:3330"  
 /clone="GQ0081.BR\_E24"  
 /sex="Hermaphrodite"  
 /tissue\_type="Vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"  
 /dev\_stage="Vascular cambium, secondary phloem and bark of trees harvested 1 and 7 days after girdling  
 /lab\_host="E. coli DH10B cells"  
 /clone\_lib="GQ008: Cambium, phloem and bark of girdled saplings"  
 /note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBlueScript II SK (+) XR; Site\_1: Eco-RI; Site\_2: Xba-I; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBlueScript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitrogen) for propagation"

## ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 817;  
 Best Local Similarity 100.0%; Pred. No. 0.69; Mismatches 0; Indels 0; Gaps 0;

RESULT 13	CO472694	CO472694	mRNA	linear	EST	09-JUL-2004
LOCUS	GQ0081.BR_1.J17	GQ0081.BR_1.J17	Cambium, phloem and bark of girdled saplings			
DEFINITION	Picea glauca cDNA clone GQ0081.BR_J17 5', mRNA sequence.					
ACCESSION	CO472694	CO472694				
VERSION	CO472694.1	GI:50141166				
KEYWORDS	EST.					
SOURCE	Picea glauca (white spruce)					
ORGANISM	Picea glauca					
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.					
JOURNAL	1 (bases 1 to 819)					
COMMENT	Morency, M.-J., Cooke, J., Pavly, N., Parsons, L., Paule, C., Sequin, A., Retzel, E., Butterfield, Y., Barber, S., Yang, G., Stott, J., Siddiqui, A., Holt, R., Marra, M. and Mackay, J.					
JOURNAL	Arboarea EST sequencing in Picea glauca (white spruce)					
CONTACT	Unpublished (2004)					
CONTACT	Contact: John Mackay					
Centre de Recherche en Biologie Forestiere						
Universite Laval						
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4						
Fax: 418 656 7493						
Email: jmackay@rsvs.ulaval.ca						
Center for Computational Genomics and Bioinformatics (CCGB), University of Minnesota, MN id Identifier: MN519649 Clone ID: GQ0081.BR_J17 Clones available through: John Mackay, Ph. D.						
professeur adjoint - Assistant professor EMAIL:						
Jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere (Forest Biology Research Center) Universite Laval Quebec, Quebec						
Plate: 1.BR row: 17 column: J						
Seq primer: M13 Reverse Primer.						
FEATURES	Location/Qualifiers					
source						
1. .819						
/organism="Picea glauca"						
/mol_type="mRNA"						
/strain="PG-653"						
/db_xref="taxon:3330"						
/clone="GQ0081.BR_J17"						
/sex="Hermaphrodite"						
/tissue_type="vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"						
/dev_stage="Vascular cambium, secondary phloem, and bark from trees harvested 1 and 7 days after girdling treatment"						
/lab_host="E. coli DH10B cells"						
/clone_lib="GQ008: Cambium, phloem and bark of girdled saplings"						
/note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBluescript II SK (+) approximately 60 cm tall; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (In vitro) for propagation"						
ORIGIN						
Query Match	100.0%	Score 25;	DB 7;	Length 819;		
Best Local Similarity	100.0%	Pred. No. 0.69;				
Matches	25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	CTGGACCCCTCTCGAGAGTCGGT 25				
Db	695	CTGGACCCCTCTCGAGAGTCGGT 719				
RESULT 15	CO472692	CO472692	mRNA	linear	EST	09-JUL-2004
LOCUS	GQ0081.BR_1.J08	GQ0081.BR_1.J08	Cambium, phloem and bark of girdled saplings			
DEFINITION	Picea glauca cDNA clone GQ0081.BR_J08 5', mRNA sequence.					
LOCUS	CO472692	CO472692				
DEFINITION	Picea glauca (white spruce)					
LOCUS	CO472674	CO472674	mRNA	linear	EST	09-JUL-2004
DEFINITION	GQ0081.BR_1.B11	GQ0081.BR_1.B11	Cambium, phloem and bark of girdled saplings			
ACCESSION	CO472674	CO472674				
VERSION	CO472674.1	GI:50141124				
KEYWORDS	EST.					
SOURCE	Picea glauca (white spruce)					
ORGANISM	Picea glauca					
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.					
AUTHORS	Morency, M.-J., Cooke, J., Pavly, N., Parsons, L., Paule, C., Sequin, A., Retzel, E., Butterfield, Y., Barber, S., Yang, G., Stott, J., Siddiqui, A., Holt, R., Marra, M. and Mackay, J.					
JOURNAL	Arboarea EST sequencing in Picea glauca (white spruce)					
COMMENT	Unpublished (2004)					
Centre de Recherche en Biologie Forestiere						
Universite Laval						
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4						
Fax: 418 656 7493						
Email: jmackay@rsvs.ulaval.ca						
Center for Computational Genomics and Bioinformatics (CCGB), University of Minnesota, MN id Identifier: MN5197207 Clone ID: GQ0081.BR_B11 Clones available through: John Mackay, Ph. D.						
professeur adjoint - Assistant professor EMAIL:						
Jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere (Forest Biology Research Center), Universite Laval Quebec, Quebec						
Plate: 1.BR row: 11 column: B						
Seq primer: M13 Reverse Primer.						
FEATURES	Location/Qualifiers					
source						
1. .820						
/organism="Picea glauca"						
/mol_type="mRNA"						
/strain="PG-653"						
/clone="GQ0081.BR_B11"						
/sex="Hermaphrodite"						
/tissue_type="vascular cambium, secondary phloem and bark of trees girdled by removing a ring of bark ca. 1 cm wide from the midpoint of the main stem"						
/db_xref="taxon:3330"						
/clone="GQ0081.BR_B11"						
/note="Organ: Main stem of 4 year old saplings, approximately 60 cm tall; Vector: pBluescript II SK (+) approximately 60 cm tall; Tissues were pooled from above and below the girdle. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (In vitro) for propagation"						
ORIGIN						
Query Match	100.0%	Score 25;	DB 7;	Length 820;		
Best Local Similarity	100.0%	Pred. No. 0.69;				
Matches	25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	CTGGACCCCTCTCGAGAGTCGGT 25				
Db	695	CTGGACCCCTCTCGAGAGTCGGT 719				

ACCESSION CO472692  
 VERSION CO472692.1  
 KEYWORD EST  
 SOURCE Picea glauca (white spruce)  
 ORGANISM Picea glauca  
 JOURNAL Eukaryota; Viridiplantae; Streptophytta; Embryophytta; Tracheophyta;  
 Spermatophytina; Coniferopsida; Coniferales; Pinaceae; Picea.  
 REFERENCE 1 (bases 1 to 825)  
 Moerency,M.-J., Cooke,J., Pavly,N., Parsons,L., Paule,C., Seguin,A.,  
 Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J.,  
 Siddiqui,A., Holt,R., Marra,M. and Mackay,J.  
 TITLE Arborea EST sequencing in Picea glauca (white spruce)  
 COMMENT Unpublished (2004)  
 Contact: John Mackay  
 Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
 Universite Laval  
 Centre de Recherche en Biologie Forestiere  
 Fax: 418 656 7793  
 Email: jmackay@ravns.ulaval.ca  
 University for Computational Genomics and Bioinformatics (CCGB),  
 University of Minnesota, MN id Identifier: MM5197621 Clone ID:  
 GQ0081.BR\_J08 Clones available through: John Mackay, Ph. D.  
 Professeur adjoint - Assistant professor EMAIL:  
 jmackay@ravns.ulaval.ca Centre de Recherche en Biologie Forestiere  
 (Forest Biology Research Center) Universite Laval Quebec, Quebec  
 CANADA G1K 7P4  
 Plate: 1.BR row: 08 column: J  
 Seq primer: M13 Reverse Primer.  
 FEATURES Location/Qualifiers  
 source  
 1. .825  
 /organism="Picea glauca"  
 /mol\_type="mRNA"  
 /strain="pg-653"  
 /db\_xref="Taxon:3330"  
 /clone="GQ0081.BR\_J08"  
 /sex="Hermaphrodite"  
 /tissue\_type="Vascular cambium, secondary phloem and bark  
 of trees girdled by removing a ring of bark ca. 1 cm wide  
 from the midpoint of the main stem"  
 /dev\_stage="Vascular cambium, secondary phloem, and bark  
 from trees harvested 1 and 7 days after girdling  
 treatment"  
 /lab\_host="E. coli DH10B cells"  
 /clone\_id="GQ0081: Cambium, phloem and bark of girdled  
 saplings"  
 /note="Organ: Main stem of 4 year old saplings,  
 approximately 60 cm tall; Vector: pBlueScript II SK (+)  
 XR; Site\_1: Eco-RI; Site\_2: Xba-I; Tissues were pooled  
 from above and below the girdle. cDNA was prepared from 5  
 mg of poly A+ selected RNA and was directionally ligated  
 into the pBlueScript II SK (+) XR vector (Stratagene),  
 transformed by electroporation into DH10B cells (In  
 vitro) for propagation"

ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 825;  
 Best Local Similarity 100.0%; P-Value: 0.69;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTGGACCCCTTCAGAGTCCCGT	25
DB	695	CTGGACCCCTTCAGAGTCCCGT	719

Search completed: March 1, 2005, 01:58:18  
 Job time : 1738.75 sec

Copyright (c) 1993 - 2005 Compugen Ltd.

### OM nucleic - nucleic search, using SW model

Run on: February 28, 2005, 23:26:41 ; Search time 192 Seconds  
 678.304 Million cell updates/sec

Title: US-10-086-062-5

Perfect score: 22

Sequence: 1 gacacgttagaatgagtcatac 22

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database : N\_Geneseq\_160sec04,\*

1: geneseqn1980s: \*  
 2: geneseqn1990s: \*  
 3: geneseqn2000s: \*  
 4: geneseqn2001ab: \*  
 5: geneseqn2001ba: \*  
 6: geneseqn2002ab: \*  
 7: geneseqn2002bs: \*  
 8: geneseqn2003ab: \*  
 9: geneseqn2003bs: \*  
 10: geneseqn2003ca: \*  
 11: geneseqn2004ab: \*  
 12: geneseqn2004bs: \*  
 13: geneseqn2004ca: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	22	100.0	22 4 AAH42764	Aab42764 A promoter element or transcription binding site.
2	22	100.0	22 6 AAD24070	Aad24070 Pea Lectin Continuation (2 of 2)
3	18.8	85.5	11000 13 ABD32868_1	Aai20672 Probe #10
4	17.2	78.2	142 4 AAI20672	Aba6716 Human fo
5	17.2	78.2	142 4 ABR65716	Aai43884 Probe #14
6	17.2	78.2	142 4 AAI45884	Aba2803 Probe #11
7	17.2	78.2	142 4 AAI11458	Aak39862 Human fo
8	17.2	78.2	142 4 AAK39862	Aak4121 Human bra
9	17.2	78.2	142 4 AAK14121	Abs3453 Human liv
10	17.2	78.2	142 4 ABS3453	Abs13960 Human gen
11	17.2	78.2	142 6 ABS13960	Aai11458 Probe #13
12	17.2	78.2	483 4 AAI11458	Aba5123 Human fo
13	17.2	78.2	483 4 ABR5123	Aai13730 Probe #14
14	17.2	78.2	483 4 AAI3730	Aba2897 Probe #13
15	17.2	78.2	483 4 ABR2897	Aak28828 Human bon
16	17.2	78.2	483 4 AAK28828	Aak01373 Human bra
17	17.2	78.2	483 4 AAB26422	Abs26422 Human liv
18	17.2	78.2	483 4 ABR01426	Abs01426 Human gen
19	17.2	78.2	483 6 ABS01426	Aab26633 Human gen
20	17.2	78.2	714 4 AAS26633	

Result No.	Score	Query Match Length	DB ID	Description
21	17.2	78.2	714 4 AAS26632	Aas6632 Human gen
22	17.2	78.2	714 4 AAK86129	Aak86129 Human imm
23	17.2	78.2	714 4 AAL36601	Aal36601 Human mus
24	17.2	78.2	714 4 AAL36602	Aal36602 Human nov
25	17.2	78.2	714 8 ABX73982	Abx73981 Human nov
26	17.2	78.2	714 8 ABX73981	Abx73981 Human nov
27	17.2	78.2	714 8 ABX59589	Abx59589 cDNA enco
28	17.2	78.2	714 8 ABX59589	Abx59589 cDNA enco
29	17.2	78.2	714 8 ADJ30339	Adj30339 Human mus
30	17.2	78.2	714 12 ADJ30340	Adj30340 Human mus
31	17.2	78.2	714 12 ADM7081_07	Continuation (8 of 8)
32	17.2	78.2	110000 11 ACM3898	Actm3898 Human gen
33	17.2	78.2	163321 11 AAQ11690	Aaq11690 Entire SE
34	16.8	76.4	6469 2 ADK59966	Adk59966 Plant DNA
35	16.2	73.6	555 10 ADQ8845	Adq8845 Novel can
36	16.2	73.6	583 13 ADR93347	Adr93347 Novel S.
37	16.2	73.6	1644 13 ADR93347	Adr93347 Novel S.
38	16.2	73.6	1980 12 ADM1890	Adm1890 S Pneumon
39	16.2	73.6	2250 10 ABX06397	Abx06397 S. pneumo
40	16.2	73.6	2253 4 AAS5670	Aas5670 Streptococ
41	16.2	73.6	2253 8 ACA49821	Aca49821 Prokaryot
42	16.2	73.6	2553 8 ADA70404	Ada70404 Rice gene
43	16.2	73.6	3234 4 ABL13054	Ab13054 Drosophil
44	16.2	73.6	3340 4 ABL19058	Abl19058 Drosophil
45	16.2	73.6	3421 4 ABL26660	Abl26660 Drosophil

### ALIGNMENTS

RESULT 1  
 AAH42764  
 ID AAH42764 standard; DNA; 22 BP.  
 XX  
 AC AAH42764;  
 XX  
 DT 01-OCT-2001 (first entry)  
 XX  
 DE A promoter element or transcription binding site.  
 KW Promoter element; transcription binding site; plant promoter; SMPER;  
 KW synthetic multimeric promoter element region; gene expression;  
 KW insect resistance; herbicide resistance; SS.  
 XX  
 OS Pisum sativum.  
 XX  
 PN WO200153476-A2.  
 PD 26-JUL-2001.  
 XX  
 PP 19-JUN-2001; 2001WO-US002024.  
 XX  
 PR 21-JAN-2000; 2000US-0177437P.  
 XX  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Bruce WB, Niu X;  
 XX  
 DR WPI; 2001-476118/51.  
 XX  
 PT New plant promoters with synthetic multimeric promoter element regions, useful in plant molecular biology, particularly in regulating gene expression in plants to increase resistance against insects or PT herbicides.  
 XX  
 PS Example 1: Fig 1; 67pp; English.  
 XX  
 CC AAH42709-72 represent promoter elements or transcription binding sites. They are used to construct synthetic multimeric promoter element regions (SMPER). The specification describes plant promoters which comprise SMPER. The plant promoters are useful in plant molecular biology, particularly in regulating gene expression in plants. The promoters are especially useful for transforming plants or plant cells.

CC e.g. to increase resistance against insects or herbicides  
 XX SQ Sequence 22 BP; 8 A; 5 C; 5 G; 4 T; 0 U; 0 Other;  
 Best Local Similarity 100.0%; Score 22; DB 4; Length 22;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACGTAGATGAGTCATCAC 22  
 ID AAD24070 standard; DNA; 22 BP.  
 Db 1 GACACGTAGATGAGTCATCAC 22

RESULT 2  
 AAD24070  
 ID AAD24070 standard; DNA; 22 BP.

XX AC AAD24070;  
 XX DT 09-APR-2002 (first entry)  
 XX DE pea lectin gene Psi element.  
 XX KW Gene expression; maize; ubiquitin heat shock element; agronomic gene; pea; lectin gene; Psi element; ds.  
 XX OS Pisum sativum.  
 PN WO200194394-A2.  
 XX PD 13-DEC-2001.  
 XX PP 08-JUN-2001; 2001WO-US018689.  
 PR 09-JUN-2000; 2000US-00590558.  
 PA (PROD-) PRODIGENE INC.  
 PT Jilka JM, Hood BE, Howard JA;  
 XX DR WPI; 2002-122117/16.  
 XX PT New promoter sequences for causing expression of a structural gene especially agronomic gene or open reading frame in a plant cell, comprises engineered versions of the maize ubiquitin promoter.  
 XX PS Claim 11; Page 55; 68pp; English.  
 CC The invention relates to a promoter sequence capable of directing expression of a nucleotide sequence in a plant cell, comprising maize ubiquitin (Ubi-1) promoter sequence with a modification so that it does not include two overlapping heat shock elements (HSE) or it directs expression to increase the endosperm/embryo expression ratio of the protein when compared to the ratio from a wild-type ubiquitin promoter. The modified Ubi-1 promoter comprises a deletion of 3', 5' or both HSEs, two non-overlapping/adjacent HSEs, replacement of HSEs with a trimer of a seed specific element from the promoter of pea lectin gene Psi<sup>1</sup> or insertion of a transcription factor binding site in the HSE region. An expression construct comprising modified Ubi-1 promoter is useful for causing expression of a structural gene (agronomic genes) or open reading frame in a plant cell. The modified Ubi-1 promoter increases expression levels beyond those observed with native ubiquitin promoter. The present invention is pea lectin gene Psi element used in the promoter of the sequence 22 BP; 8 A; 5 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 6; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 0.24; Mismatches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACGTAGATGAGTCATCAC 22  
 CC ||||||| 22

Db 1 GACACGTAGATGAGTCATCAC 22

RESULT 3  
 ABD32968 1  
 Continuation (2 of 8) of ABD32968 from base 10001 (Human cancer-associated genomic DNA);  
 Sequence split into 8 fragments LOCUS ABD32968 Accession Abd32968  
 WP Fragment Name Begin End  
 WP ABD32968\_0 1 110000  
 WP ABD32968\_1 100001 210000  
 WP ABD32968\_2 200001 310000  
 WP ABD32968\_3 300001 410000  
 WP ABD32968\_4 400001 510000  
 WP ABD32968\_5 500001 610000  
 WP ABD32968\_6 600001 710000  
 WP ABD32968\_7 700001 779603

Query Match 85.5%; Score 18.8; DB 13; Length 110000;  
 Best Local Similarity 90.9%; Pred. No. 29; Mismatches 20; Indels 0; Gaps 0;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GACACGTAGATGAGTCATCAC 22  
 Db 42953 GACACGTAGATGAGTCATCAC 42974

RESULT 4  
 AAI20672  
 ID AAI20672 standard; DNA; 142 BP.  
 XX AC AAI20672;  
 XX DT 12-OCT-2001 (first entry)  
 DE Probe #10605 for gene expression analysis in human cervical cell sample.  
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer; ss.  
 XX OS Homo sapiens.  
 PN WO200157278-A2.  
 XX PD 09-AUG-2001.  
 XX PP 30-JAN-2001; 2001WO-US000670.  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-0308403.  
 PR 03-AUG-2000; 2000US-0032365.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000US-0024263.  
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX DR WPI; 2001-48890153.  
 XX PS Claim 25; SEQ ID NO 10605; 487pp; English.

CC Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human cervical epithelial cells.  
 XX SQ Sequence 22 BP; 8 A; 5 C; 5 G; 4 T; 0 U; 0 Other;  
 CC The present invention relates to human single exon nucleic acid probes (SEMP). The present sequence is one such probe. The SEMP are derived from human HeLa cells. The SEMP can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer. Note: The sequence data

CC for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at  
 CC AAI5884  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;  
 Query Match 78.2%; Score 17.2; DB 4; Length 142;  
 Best Local Similarity 86.4%; Pred. No. 86;  
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 GACAGCTAGATGAGTCATCAC 22  
 Db 80 GACACATAAAATCAGTCATCAC 101  
 OS Homo sapiens.  
 XX  
 RESULT 5  
 ABA65716  
 ID ABA65716 standard; DNA; 142 BP.  
 XX  
 AC ABA65716;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Human foetal liver single exon nucleic acid probe #14021.  
 XX  
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 AC 30-JAN-2001; 2001WO-US000669.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-0207456P.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0234687P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.  
 XX  
 PS Claim 25; SEQ ID NO 14570; 654pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SNPs).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders  
 XX  
 SQ Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;  
 Query Match 78.2%; Score 17.2; DB 4; Length 142;  
 Best Local Similarity 86.4%; Pred. No. 86;  
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 GACAGCTAGATGAGTCATCAC 22  
 Db 80 GACACATAAAATCAGTCATCAC 101  
 OS Homo sapiens.  
 XX  
 RESULT 6  
 ABA65884  
 ID AAI45884 standard; DNA; 142 BP.  
 XX  
 AC AAI45884;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Probe #14570 used to measure gene expression in human placenta sample.  
 XX  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder; SB.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 AC 30-JAN-2001; 2001WO-US000663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0234687P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.  
 XX  
 PS Claim 25; SEQ ID NO 14570; 654pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SNPs).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders  
 XX  
 SQ Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;  
 Query Match 78.2%; Score 17.2; DB 4; Length 142;  
 Best Local Similarity 86.4%; Pred. No. 86;  
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 GACAGCTAGATGAGTCATCAC 22  
 Db 80 GACACATAAAATCAGTCATCAC 101  
 OS Homo sapiens.  
 XX  
 RESULT 7  
 ABA22803  
 ID ABA22803 standard; DNA; 142 BP.  
 XX  
 AC ABA32803;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Probe #11269 for gene expression analysis in human heart cell sample.  
 XX  
 KW Human; gene expression; heart; microarray; vascular system; probe;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease; ss.  
 XX  
 OS Homo sapiens.

XX  
PN WO200157274-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000566.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-06059408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000US-00024253.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PR Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
PR XX  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PR Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
PR XX  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts.  
XX  
RS Claim 4; SEQ ID NO 11269; 530bp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes for  
measuring human gene expression in a sample derived from human heart. The  
present sequence is one such probe. The probes may be used for  
predicting, measuring and displaying gene expression in samples derived  
from the human heart via microarrays. By measuring gene expression, the  
probes are useful for predicting, diagnosing, grading, staging,  
monitoring and prognosis diseases of the human heart and vascular system  
e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
congenital heart disease. Note: The sequence data for this patent did not  
form part of the printed specification, but was obtained in electronic  
format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)  
XX  
SQ Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;  
Query Match 78.2%; Score 17.2; DB 4; Length 142;  
Best Local Similarity 86.4%; Pred. No. 86;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GACACGTAGATGACTCATCAC 22  
Db 80 GACACATAAAATCAGTCATCAC 101  
RESULT 9  
ID AAK14121  
ID AAK14121 standard; DNA; 142 BP.  
XX  
AC AAK14121;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DB Human brain expressed single exon probe SEQ ID NO: 14112.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000667.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-0008408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000US-00024253.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PT Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
PR WPI; 2001-488900/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human bone marrow.  
XX  
PR Example 4; SEQ ID NO 14419; 658pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
probes which are derived from genomic sequences expressed in the human  
bone marrow. They can be used to measure gene expression in bone marrow  
samples, which may enable the improved diagnosis and treatment of cancers  
such as lymphoma, leukaemia and myeloma. The present sequence is one of  
the probes of the invention.  
XX  
Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;  
Query Match 78.2%; Score 17.2; DB 4; Length 142;  
Best Local Similarity 86.4%; Pred. No. 86;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GACACGTAGATGACTCATCAC 22  
Db 80 GACACATAAAATCAGTCATCAC 101  
RESULT 9  
ID AAK14121  
ID AAK14121 standard; DNA; 142 BP.  
XX  
AC AAK14121;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DB Human brain expressed single exon probe SEQ ID NO: 14112.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000667.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-0008408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000US-00024253.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PT Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
PR WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains.

PS Example 4; SEQ ID NO 14112; 650pp + Sequence Listing; English.

XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
which may enable the diagnosis and improved treatment of nervous system  
diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
epilepsy and cancers. The present sequence is one of the probes of the  
CC invention  
XX  
SQ sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 142;  
Best Local Similarity 86.4%; Pred. No. 86;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACGAGAAGTGTATCAC 22  
Db 80 GACACATAAACAGTCATCAC 101

RESULT 10

ABS39453  
ID ABS39453 standard; DNA; 142 BP.  
XX  
AC ABS39453;  
XX  
DT 25-FEB-2003 (first entry)

DE Human liver single exon probe, SEQ ID NO 14443.  
XX Human; single exon nucleic acid probe; liver; cirrhosis;  
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;  
KW coronary heart disease; ss.  
XX OS Homo sapiens.  
XX PN WO200157273-A2.  
XX PD 09-AUG-2001..  
XX PR 30-JAN-2001; 2001WO-US000664.  
XX PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-488898/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.  
XX PS Claim 4; SEQ ID NO 14443; 650pp; English.  
XX  
CC The invention relates to a single exon nucleic acid probe (SNP) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridizes at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (II) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinæmia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABS39453 represent human  
CC liver single exon nucleic acid probes of the invention. Note: The

CC sequence information for this patent does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;  
Query Match 78.2%; Score 17.2; DB 4; Length 142;  
Best Local Similarity 86.4%; Pred. No. 86;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACACGAGAAGTGTATCAC 22  
Db 80 GACACATAAACAGTCATCAC 101

RESULT 11

ABS13960  
ID ABS13960 standard; DNA; 142 BP.  
XX  
AC ABS13960;  
XX  
DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe ORF from lung SBO ID No 13951.  
XX Human; ds; single exon probe; asthma; lung cancer; COPD; IUD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease; open reading frame; ORF.  
XX OS Homo sapiens.  
XX PN WO200186003-A2.  
XX PD 15-NOV-2001..  
XX PR 30-JAN-2001; 2001WO-US000665.  
XX PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2002-114103/15.

XX PT Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples.  
XX PS Claim 4; SEQ ID NO 13951; 63pp; English.

CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of probes  
CC ; the novel set of probes which hybridise at high stringency to a nucleic  
CC acid expressed in the human lung; measuring gene expression in a sample  
CC derived from human lung, comprising (a) contacting the array with a  
CC collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of the

array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridization of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridization to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probe/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberculous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemangioma, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagener syndrome, fibrocytic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a single exon probe open reading frame of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://ftp.wipo.int/pub/published\\_pct\\_sequences](ftp://ftp.wipo.int/pub/published_pct_sequences)

SQ Sequence 142 BP; 52 A; 40 C; 20 G; 30 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 6; Length 142;  
Best Local Similarity 86.4%; Pred. No. 86; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACACCTAGATGAGTCATCC 22  
Db 80 GACACATAAACATCAGTCAC 101

RESULT 12

AAT11458 ID AAT11458 standard; DNA; 483 BP.  
AC AAT11458;  
XX DT 12-OCT-2001 (first entry)

DE Probe #1391 for gene expression analysis in human cervical epithelial cell sample.

KW Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer; SS;  
XX OS Homo sapiens.

PN WO200157277-A2.  
XX PD 09-AUG-2001.

PP 30-JAN-2001; 2001WO-US000669.

PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

array; identifying exons in a eukaryotic genome, comprising (a)

algorithmsically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridization of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene,

comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridization to a single exon microarray having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one

of 12011 sequences, mentioned in the specification, or encoded by the probe/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,

tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemangioma, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagener syndrome, fibrocytic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a single exon probe open reading frame of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://ftp.wipo.int/pub/published\\_pct\\_sequences](ftp://ftp.wipo.int/pub/published_pct_sequences)

XX

Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human cervical epithelial cells.

CC Claim 25; SEQ ID NO 1391; 487pp; English.

The present invention relates to human single exon nucleic acid probes (SNPs). The present sequence is one such probe. The SNPs are derived from human HeLa cells. The SNPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://ftp.wipo.int/pub/published\\_pct\\_sequences](ftp://ftp.wipo.int/pub/published_pct_sequences)

XX

Sequence 483 BP; 151 A; 106 C; 99 G; 127 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 483;  
Best Local Similarity 86.4%; Pred. No. 1e+02; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACAGCTGAACTGACTTCAC 22  
Db 402 GACACATAAACATCAGTCAC 423

RESULT 13

ABA53123 ID ABA53123 standard; DNA; 483 BP.  
AC ABA53123;

XX DT 01-FEB-2002 (first entry)

Human foetal liver single exon nucleic acid probe #1428.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; SS.

OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

PP 30-JAN-2001; 2001WO-US000669.

PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR XX WPI; 2001-483447/52.

PR 04-FEB-2000; 2000US-020745P.  
PR 30-JUN-2000; 2000US-0060840B.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-

CC specification, but was obtained in electronic format directly from WIPO  
 at [ftp.wipo.int/pub/published\\_pct\\_sequences](ftp://ftp.wipo.int/pub/published_pct_sequences)

XX Sequence 483 BP; 151 A; 106 C; 99 G; 127 T; 0 U; 0 Other;

SQ Best Local Similarity 86.4%; Pred. No. 1e+02; Length 483;  
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACACTTAGATGAGTCATCA 22

Db 402 GACACATAAAATCAGTCATCA 423

RESULT 14

AAI32730 ID AAI32730 standard; DNA; 483 BP.

XX AC AAI32730;

XX DT 17-OCT-2001 (first entry)

XX DE Probe #1416 used to measure gene expression in human Placenta sample.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KW genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO20015272-A2.

XX PD 09-AUG-2001.

XX PR 30-JAN-2001; 2001WO-US000666.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00523365.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-023359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488899/53.

XX PT Single exon nucleic acid probes for analyzing gene expression in human hearts.

XX PR Claim 1; SEQ ID NO 1363; 530PP; English.

XX PS

CC The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosis diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](ftp://ftp.wipo.int/pub/published_pct_sequences)

XX Sequence 483 BP; 151 A; 106 C; 99 G; 127 T; 0 U; 0 Other;

XX CC The present invention relates to single exon nucleic acid probes (SENP). The present sequence is one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta.

XX PS Claim 25; SEQ ID NO 1416; 654PP; English.

XX CC Query Match 78.2%; Score 17.2; DB 4; Length 483;

CC PT Best Local Similarity 86.4%; Pred. No. 1e+02; Length 483;  
 Matches 19; Conservative 0; Mismatched 3; Indels 0; Gaps 0;

QY 1 GACACTTAGATGAGTCATCA 22

Db 402 GACACATAAAATCAGTCATCA 423

Search completed: February 28, 2005, 23:51:50  
 Job time : 198 secs

*This Page Blank (uspto)*



**ORGANISM** Agrobacterium tumefaciens  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

**REFERENCE** 1  
**AUTHORS** Jilka, J.M., Hood, E.B. and Howard, J.A.  
**TITLE** Novel plant promoter sequences and methods of use for same  
**JOURNAL** Patent: WO 014394-A 5 13-DEC-2001;  
**FEATURES** source  
1. .22  
/organism="Agrobacterium tumefaciens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:358"

**ORIGIN**

Query Match Best Local Similarity 100 %; Score 22; DB 6; Length 22;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 1 GACACCTGAGATGAGTCATCAC 22  
**Db** 1 GACACCTGAGATGAGTCATCAC 22

**RESULT 3**  
**AY192160** AY192160 synthetic construct multi-copy enhancer promoter element, complete  
**LOCUS** AY192160  
**DEFINITION** SYNTHETIC CONSTRUCT  
**SEQUENCE** sequence.  
**VERSION** AY192160.1  
**KEYWORDS** synthetic construct  
**SOURCE** synthetic construct  
**ORGANISM** Other sequences; artificial sequences.  
**REFERENCE** 1  
**AUTHORS** Yao, Q., Peng, R. and Xiong, A.  
**TITLE** An artificial promoter containing many copies of enhancers can increase phytase expression in transgenic rape  
**JOURNAL** Unpublished  
**REFERENCE** 2  
**AUTHORS** Yao, Q., Peng, R. and Xiong, A.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (07-DEC-2002) Biotechnology Research Center, Shanghai Academy of Agriculture Science, Beidi Road 2901, Shanghai, Shanghai 201106, China  
**FEATURES** source  
1. .827  
/organism="synthetic construct"  
/mol\_type="Genomic DNA"  
/db\_xref="taxon:32630"

**RESULT 4**  
**PBAEBC** PBAEBC  
**LOCUS** PBAEBC  
**DEFINITION** Pea PsII gene encoding lectin, complete cds.  
**ACCESSION** M18160  
**VERSION** M18160.1  
**KEYWORDS** lectin.  
**SOURCE** Pisum sativum (pea)  
**ORGANISM** Pisum sativum (pea)  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids I; Fabales; Papilionoideae; Viciae; Pisum.  
**REFERENCE** 1  
**AUTHORS** Kaminski, P.A., Buffard, D. and Strosberg, A.D.  
**TITLE** The pea lectin gene family contains only one functional gene  
**JOURNAL** Plant Mol. Biol. 9, 497-507 (1987)  
**COMMENT** Original source text: Pea seed, cDNA to mRNA, clone lambda-I-101.  
**FEATURES** source  
1. .1659  
/organism="Pisum sativum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3888"  
<1. .1659  
/product="lectin mRNA"  
463. .1280  
/note="lectin"  
/codon\_start=1  
/protein\_id="AA33676.1"  
/translation="MAISIQOMISFYATFLSTLTTILFFKVNSTETTSFLITKFSPODQNLIFQDGTYTKEKLTKTAKNTVGRALLSPHINWDRTGIVNANFVTDKTOVAVERDFTAPUDTFGGVILGVNSAEDVKTDTQVAVRSLWWSFHESLSGTSSSKQADA" SNRHIGIDVWNSTKSIVTKSIVLQLNGEANNVIAFNAINTLIVSUTPSLNEEV TSYTLSDVVSLSKDVPKVNPWRIGFSATTGAEYAHAEVULSWWSFHESLSGTSSSKQADA"

**ORIGIN**

Query Match Best Local Similarity 100 %; Score 22; DB 8; Length 1659;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 1 GACACCTGAGATGAGTCATCAC 22  
**Db** 392 GACACCTGAGATGAGTCATCAC 413

**RESULT 5**  
**A14134** A14134 promoter and lectin gene.  
**LOCUS** A14134  
**DEFINITION** promoter and lectin gene.  
**ACCESSION** A14134  
**VERSION** A14134.1  
**KEYWORDS** GT:490034  
**SOURCE** Pisum sativum (pea)  
**ORGANISM** Pisum sativum (pea)  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilioideae; Viciae; Pisum.  
**REFERENCE** 1  
**AUTHORS** Hepher, A., Edwards, G.A. and Gatehouse, J.A.  
**TITLE** Improvements relating to transgenic plants  
**JOURNAL** Patent: EP 0351924-A 1-24-JAN-1990;  
SIBYL INTERNATIONALE RESEARCH MAATSCHAPPIJ B.V.; Nickerson Seeds Limited  
**FEATURES** source  
1. .1701  
/organism="Pisum sativum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3888"

**ORIGIN**

Query Match Best Local Similarity 100 %; Score 22; DB 12; Length 827;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 1 GACACCTGAGATGAGTCATCAC 22

gene  
473. .1300  
/gene="leca"  
CDS  
473. .1300  
/gene="leca"  
/codon\_start=1  
/product="lectin"  
/protein\_id="CAA01149.1"  
/db\_xref="GI:490035"  
/translation="MASTIOTOMISFYAIFSLTILTFKVNSTETPSLITKSPD  
QNLIFQDGTYTKELITKAKNTVGRALNNSPPIHWDREGVANFVSTFFVIN  
APMSYVNAUDGFTPIPAVDTKPGGGVGFVNSAEYKTTGVAEVDTFTNAWDP  
SNDRHGIDVNSISVNTKSWKLONGERANVUATANVLTYSITPNSLEEN  
TSTYLSDVSLKDVPEWURIGFSATTGABYAAHEVLWSWSFHSELSGISSSKOADA"

ORIGIN  
Query Match 100.0%; Score 22; DB 6; length 1701;  
Best Local Similarity 100.0%; Pred. No. 0.89; 0; Gaps 0;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 414 GACACGTAGATGAGTCATCAC 435

RESULT 6  
PSLCA  
LOCUS pslCA  
DEFINITION pslum sativum leca gene for lectin.  
ACCESSION Y00440  
VERSION 1  
KEYWORDS leca gene; lectin.  
ORGANISM pslum sativum (pta)  
REFERENCE 1 (bases 1 to 1701)  
AUTHORS Gatehouse,J.A., Bown,D., Evans,I.M., Gatehouse,L.N., Jobes,D.,  
preston,P. and Croy,R.R.  
TITLE Sequence of the seed lectin gene from pea (*Psium sativum* L.)  
JOURNAL Nucleic Acids Res. 15 (18), 7642 (1987)  
MEDLINE 88015625  
PUBLMED 3658708  
REFERENCES 2 (bases 1 to 1701)  
AUTHORS Gatehouse,J.A.  
TITLE Direct Submission  
JOURNAL Submitted (07-SEP-1987) Gatehouse,J.A., Botany Department,  
University of Durham, South Road, Durham, DH1 3LE, England  
FEATURES Source  
B\_SOURCE Location,Qualifiers

1. .1701  
/organism="Psium sativum"  
/mol\_type="genomic DNA"  
/strain="Feldham First"  
/db\_xref="taxon:3888"  
/tbl\_Bioe\_type="Beed"  
/note="TATA-Box"  
473. .1300  
/codon\_start=1  
/product="lectin-precursor" (AA -30 to 245)"  
/protein\_id="CAA66497.1"  
/db\_xref="GI:490035"  
/db\_xref="UniProt:SWISS-PROT: P02867"  
/translation="MASIQTOMISFYAIFSLTILTFKVNSTETPSLITKSPD  
QNLIFQDGTYTKELITKAKNTVGRALNNSPPIHWDREGVANFVSTFFVIN  
APMSYVNAUDGFTPIPAVDTKPGGGVGFVNSAEYKTTGVAEVDTFTNAWDP  
SNDRHGIDVNSISVNTKSWKLONGERANVUATANVLTYSITPNSLEEN  
TSTYLSDVSLKDVPEWURIGFSATTGABYAAHEVLWSWSFHSELSGISSSKOADA"  
/note="signal peptide (AA -30 to -1)"  
mat\_peptide  
563. .1297

ORIGIN  
Query Match 100.0%; Score 22; DB 6; length 1701;  
Best Local Similarity 100.0%; Pred. No. 0.89; 0; Gaps 0;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 414 GACACGTAGATGAGTCATCAC 435

RESULT 7  
PSLLECTIN  
LOCUS pslLECTIN  
DEFINITION P sativum psl gene for Psl lectin.  
ACCESSION X66368  
VERSION X66368.1  
KEYWORDS lectin; psl gene; Psl lectin.  
ORGANISM psium sativum (pea)  
REFERENCE 1 (bases 1 to 3360)  
AUTHORS de Pater,S., Pham,K., Chua,N.H., Memelink,J. and Kijne,J.  
TITLE A 22-bp fragment of the pea lectin promoter containing essential  
TGAC-like motifs confers seed-specific gene expression  
JOURNAL Plant Cell 5 (8), 877-886 (1993)  
MEDLINE 94004020  
PUBLMED 8400870  
REFERENCES 2 (bases 1 to 1713)  
AUTHORS Hoedemaeker,F.J., Richardson,M., Diaz,C.L., de Pater,B.S. and  
Kijne,J.W.  
TITLE Pea (*Psium sativum* L.) seed islectins 1 and 2 and pea root lectin  
result from carboxypeptidase-like processing of a single gene  
product  
JOURNAL Plant Mol. Biol. 24 (1), 75-81 (1994)  
MEDLINE 94154245  
PUBLMED 8411028  
REFERENCES 3 (bases 1 to 3360)  
AUTHORS de Pater,S., Katagiri,F., Kijne,J. and Chua,N.H.  
TITLE bZIP proteins bind to a palindromic sequence without an ACGT core  
located in a seed-specific element of the pea lectin promoter  
JOURNAL Plant J. 6 (2), 133-140 (1994)  
MEDLINE 95004657  
PUBLMED 7920707  
REFERENCES 4  
AUTHORS de Pater,B.S.  
TITLE Direct Submission  
JOURNAL Submitted (19-MAY-1992) B.-S. de Pater, Center for Phytotechnology,  
Botanical Laboratory, Nonnensteeg 3, 2311 VJ Leiden, THE  
NETHERLANDS  
revised by [3] MAT  
5 (bases 1 to 3360)  
de Pater,B.S.  
TITLE Direct Submission  
JOURNAL Submitted (27-OCT-1994) B.-S. de Pater, Center for Phytotechnology,  
Clusius Laboratorium, Wassenaarseweg 64, 2333 AL Leiden, THE  
NETHERLANDS  
COMMENT On Nov 1, 1994 this sequence version replaced gi:20803.  
FEATURES source  
/organism="Psium sativum"  
/mol\_type="genomic DNA"  
/variety="Feldham First"

**ORIGIN**  
**RESULT 8**  
**AC01143**  
**LOCUS** AC01143. DEFINITION Homo sapiens chromosome 18, clone RP11-1015, complete sequence.  
**VERSION** AC01143.6  
**KEYWORDS** RTG.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
**REFERENCE** 1 (bases 1 to 136492)  
**AUTHORS** Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
**TITLE** Human chromosomes 18, clone RP11-1015  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 136492)  
**AUTHORS** Birren,B., Linton,L., Nusbaum,C., Kann,L., Landier,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckery,R., Boguslavskiy,L., Bouknighter,B., Brown,A., Castie,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearlano,K., Devar,K., Domino,M., Donegan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galigan,J., Gardina,S., Grant,G., Hacos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lebovitz,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEvon,P., McCork,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Navlak,J., Norman,C.H., O'Connor,T., O'Connell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Teafay,S., Tirrell,A., Vasilev,H., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.  
**JOURNAL** Direct Submission  
**REFERENCE** Submitted (01-CH-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**AUTHORS** Barren,B., Linton,L., Nusbaum,C., Landier,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Bouknighter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chooper,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearlano,K., Devar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,  
**COMMENT** On Mar 28, 2001 this sequence version replaced gi:13324843.  
**TITLE** Submitted (02-APR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**JOURNAL** All repeats were identified using repeatMasker:  
Smit,A.F.A. & Green,P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
**FEATURES**  
**source**  
**repeat\_region** /organism="Homo sapiens"  
/biol\_type="Genomic DNA"  
/db\_xref="taxon:9606"  
/map="18"  
/clone="RP11-1015"  
/clone\_libr="RPCI-11 Human Male BAC"  
/rpt\_family="MER39"  
/complement(1511..1896)  
/repeat\_region /rpt\_family="THE1C"  
/complement(1897..1983)  
/rpt\_family="MER39"  
/complement(2259..2291)  
/rpt\_family="L2"  
/complement(250..1510)  
/repeat\_region /rpt\_family="THE1B"  
/complement(3249..3402)  
/repeat\_region /rpt\_family="Aluy"  
/complement(3657..3754)  
/rpt\_family="Alujo"  
/complement(3956..3985)  
/repeat\_region /rpt\_family="AT\_rich"  
/complement(4007..4295)  
/repeat\_region /rpt\_family="Alujo"  
/complement(5578..5587)  
/rpt\_family="L1MB"  
/complement(5827..5887)  
/repeat\_region /rpt\_family="L1MB4a"  
/complement(5836..6242)  
/rpt\_family="MSTB1"  
/complement(6481..6807)  
/rpt\_family="YALuQ"  
/complement(7221..7394)  
/repeat\_region /rpt\_family="MTR"  
/complement(7429..8364)  
/rpt\_family="LTRB"  
/complement(10152..10187)  
/rpt\_family="(TG)n"

repeat_region	complement(10614..11362)	/rpt_family="LIMC4a"
repeat_region	/rpt_family="L2"	/rpt_family="LIMA4" . 38509
repeat_region	11364..11700	/rpt_family="L2"
repeat_region	11903..12038	/rpt_family="L2"
repeat_region	12793..13820	/rpt_family="L2"
repeat_region	/rpt_family="AT_rich"	/rpt_family="AT_rich"
repeat_region	complement(13170..13201)	complement(13170..13201)
repeat_region	/rpt_family="U2"	/rpt_family="U2"
repeat_region	complement(13287..13340)	complement(13287..13340)
repeat_region	/rpt_family="MER5A"	/rpt_family="MER5A"
repeat_region	13395..14443	13395..14443
repeat_region	/rpt_family="MER5A"	/rpt_family="MER5A"
repeat_region	complement(13401..13442)	complement(13401..13442)
repeat_region	/rpt_family="MER5B"	/rpt_family="MER5B"
repeat_region	14965..15007	14965..15007
repeat_region	/rpt_family="TCTA"	/rpt_family="TCTA"
repeat_region	15008..15294	15008..15294
repeat_region	/rpt_family="ALusX"	/rpt_family="ALusX"
repeat_region	15296..15317	15296..15317
repeat_region	/rpt_family="AT_rich"	/rpt_family="AT_rich"
repeat_region	15572..15759	15572..15759
repeat_region	/rpt_family="MER5A"	/rpt_family="MER5A"
repeat_region	complement(16050..16156)	complement(16050..16156)
repeat_region	/rpt_family="LTR1"	/rpt_family="LTR1"
repeat_region	complement(16219..16611)	complement(16219..16611)
repeat_region	17548..18358	17548..18358
repeat_region	/rpt_family="HUERS-P2"	/rpt_family="HUERS-P2"
repeat_region	complement(16612..17391)	complement(16612..17391)
repeat_region	/rpt_family="MLT1C"	/rpt_family="MLT1C"
repeat_region	19131..19573	19131..19573
repeat_region	/rpt_family="HALL"	/rpt_family="HALL"
repeat_region	19554..19621	19554..19621
repeat_region	/rpt_family="HALL"	/rpt_family="HALL"
repeat_region	19818..19845	19818..19845
repeat_region	/rpt_family="AT_rich"	/rpt_family="AT_rich"
repeat_region	complement(20169..20627)	complement(20169..20627)
repeat_region	/rpt_family="LIM4"	/rpt_family="LIM4"
repeat_region	complement(20628..21367)	complement(20628..21367)
repeat_region	/rpt_family="LIPAT7"	/rpt_family="LIPAT7"
repeat_region	complement(21368..21519)	complement(21368..21519)
repeat_region	/rpt_family="LIM4"	/rpt_family="LIM4"
repeat_region	complement(21872..22228)	complement(21872..22228)
repeat_region	/rpt_family="MER39"	/rpt_family="MER39"
repeat_region	complement(22232..22327)	complement(22232..22327)
repeat_region	/rpt_family="LTP48B"	/rpt_family="LTP48B"
repeat_region	complement(22378..22857)	complement(22378..22857)
repeat_region	/rpt_family="MER39"	/rpt_family="MER39"
repeat_region	24790..24827	24790..24827
repeat_region	/rpt_family="(T)n"	/rpt_family="(T)n"
repeat_region	25137..25201	25137..25201
repeat_region	/rpt_family="(TATATG)n"	/rpt_family="(TATATG)n"
repeat_region	complement(25277..26219)	complement(25277..26219)
repeat_region	/rpt_family="LIM4"	/rpt_family="LIM4"
repeat_region	complement(26222..26375)	complement(26222..26375)
repeat_region	/rpt_family="LTP22"	/rpt_family="LTP22"
repeat_region	complement(26404..26909)	complement(26404..26909)
repeat_region	/rpt_family="LTP9"	/rpt_family="LTP9"
repeat_region	complement(26911..34208)	complement(26911..34208)
repeat_region	/rpt_family="HUERS-P3b"	/rpt_family="HUERS-P3b"
repeat_region	34209..34324	34209..34324
repeat_region	/rpt_family="CATTCn"	(bases 1 to 19556)
repeat_region	complement(34325..34399)	(bases 1 to 19556)
repeat_region	/rpt_family="HUERS-P2b"	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barnai,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Dodge,S., Faro,S., Ferreiro,P., FitzHugh,W., Gage,D., Galagan,J., Gardyne,S., Gindele,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Rulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Larocque,K., Lamazares,R., Landers,T., Lehozcky,J., Levine,R., Liu,J., MacLean,C., McDonald,P., Marguia,N., Matthews,C., McCarthy,M., Mchowa,P., McKernan,K., McPheters,R., Meldrim,J., Meneus,L., Mihowa,T., Mlenga,V., Murphy,T., Nayor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhangsi,P., Pierre,N., Pollard,V., Raymond,C., Retta,R., Ribbeck,M., Riley,R., Rose,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Sougnec,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strausberg,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Triglio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,A., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.
repeat_region	Direct Submission	Submitted (07-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3	3
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barnai,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Dodge,S., Faro,S., Ferreiro,P., FitzHugh,W., Gage,D., Galagan,J.,	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barnai,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Dodge,S., Faro,S., Ferreiro,P., FitzHugh,W., Gage,D., Galagan,J.,

Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyne,S.,  
 Giude,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Haford,A., Horton,L., Huime,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Katsas,A., Kellis,J., Levine,R., Liu,G.,  
 Lamazares,R., Landers,T., Lethoczy,J., Levin,R., Liu,G.,  
 Maclean,C., McDonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrim,J.,  
 Meneus,L., Mihova,T., Milenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Nobru,C., Norman,C.H., O'Connell,T., O'Donnell,P., O'Neil,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Rettig,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rotetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,  
 Seaman,S., Severy,P., Spender,B., Strange-Thomann,N., Stofanovic,N.,  
 Strauss,N., Subramanian,A., Talman,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,J., Trigilio,J., Vassilier,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.  
 Direct Submission  
 JOURNAL Submitted (28-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 01141, USA  
 COMMENT On Nov 22, 2001 this sequence version replaced gi:14192992.  
 All repeats were identified using RepeatMasker:  
 Smit, A.P.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 Center project name: L12340  
 Center clone name: 869\_0\_1  
 -----  
**FEATURES**  
 source  
 -----  
 Location/Qualifiers  
 1. .-19156  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="Taxon:9606"  
 /chromosome="18"  
 /map="18"  
 /clone="RP11-86901"  
 /clone lib="RPCI-11 Human Male BAC"  
 123. .158  
 /rpt\_family="(TG)n"  
 unsure  
 /note=<30 qual SNGL region"  
 492. .529  
 unsure  
 /note=<30 qual SNGL region"  
 539. .575  
 unsure  
 /note=<30 qual SNGL region"  
 582. .590  
 unsure  
 /note=<30 qual SNGL region"  
 repeat\_region complement(585..1333)  
 /rpt\_family="LIMC/D"  
 602. .612  
 unsure  
 /note=<30 qual SNGL region"  
 710. .714  
 unsure  
 /note=<30 qual SNGL region"  
 1202. .1207  
 unsure  
 /note=<30 qual SNGL region"  
 1218. .1224  
 unsure  
 /note=<30 qual SNGL region"  
 1335. .1671  
 repeat\_region /rpt\_family="L2"  
 repeat\_region 1874. .2009  
 repeat\_region /rpt\_family="L2"  
 repeat\_region 2764. .2791  
 repeat\_region /rpt\_family="AT-rich"  
 repeat\_region complement(3141..3172)  
 repeat\_region /rpt\_family="L2"  
 repeat\_region complement(3256..3311)  
 repeat\_region /rpt\_family="MER5A"  
 repeat\_region 3366..3414  
 repeat\_region /rpt\_family="MER5A"  
 repeat\_region complement(3372..3413)  
 -----  
 repeat\_region /rpt\_family="MER5B"  
 4936. .4978  
 repeat\_region /rpt\_family="(TCTA)n"  
 complement(4979..5265)  
 /rpt\_family="AluSX"  
 repeat\_region 5267. .5288  
 /rpt\_family="AT-rich"  
 5543. .5730  
 /rpt\_family="MER5A"  
 complement(6021..6127)  
 /rpt\_family="LTR1"  
 complement(6190..6582)  
 /rpt\_family="HUBS\_P2"  
 complement(6583..7362)  
 /rpt\_family="LTR1"  
 /rpt\_family="HALL"  
 7519..8129  
 /rpt\_family="HALL"  
 8457..8618  
 /rpt\_family="MER96B"  
 complement(8625..9073)  
 /rpt\_family="MLTC"  
 9102..9144  
 /rpt\_family="HALL"  
 9525..9532  
 /rpt\_family="HALL"  
 9789..9816  
 /rpt\_family="AT-rich"  
 complement(10140..10598)  
 /rpt\_family="LIM4"  
 complement(11339..11338)  
 /rpt\_family="LPA7"  
 complement(11339..11490)  
 /rpt\_family="LIM4"  
 complement(11843..12199)  
 /rpt\_family="LIM4"  
 complement(12203..12298)  
 /rpt\_family="LTP48B"  
 complement(12349..12828)  
 /rpt\_family="MER39"  
 14761..14797  
 /rpt\_family="(T)n"  
 15101..15171  
 /rpt\_family="(TATAG)n"  
 complement(15247..16189)  
 /rpt\_family="LIMA4"  
 complement(1619..16345)  
 /rpt\_family="LTR9"  
 complement(16374..16779)  
 /rpt\_family="LTR9"  
 complement(16681..24179)  
 /rpt\_family="HUEBS\_P30"  
 complement(22065..22669)  
 /rpt\_family="LTR9"  
 complement(22674..24918)  
 /note=<30 qual SNGL region"  
 24180..24295  
 /rpt\_family="(CATTC)n"  
 complement(24296..24370)  
 /rpt\_family="LIMA4-P30"  
 complement(24400..24918)  
 /rpt\_family="LTR9"  
 complement(24923..27349)  
 /rpt\_family="LIMA4"  
 complement(27404..28480)  
 /rpt\_family="LIMA4"  
 complement(28507..29407)  
 /rpt\_family="LIMA4"  
 complement(29803..29940)  
 /rpt\_family="FLAM\_C"  
 complement(30195..30485)  
 /rpt\_family="AluSq"  
 complement(310868..30991)  
 /rpt\_family="MIR"  
 repeat\_region 31347..31479





----- Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
\*(see [http://www.hgic.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgic.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as run of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 224161: contig of 224161 bp in length

\* 224162 224261: gap of unknown length

\* 224262: contig of 1367 bp in length

\* 225629 225728: gap of unknown length

\* 225729 226810: contig of 1082 bp in length.

FEATURES Source Location/Qualifiers

1. . 226810 /organism="Rattus norvegicus"

/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"

/clone="CH230-81B"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

misc\_feature 1. . 1284 /note="wgb contig"

223761. . 224161 /note="wgb contig"

-----

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more; than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Swi: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep). Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhifeng Bao and Sean Eddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see [http://www.sanger.ac.uk/Projects/D\\_rerio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml)

CH211-279M15 is from a CHORI-211 BAC library

VECTOR: pTRABAC2.1

FEATURES Source Location/Qualifiers

1. . 146407 /organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="CH211-279M15"  
/clone\_lib="CHORI-211"

ORIGIN

Query Match 81.8%; Score 18; DB 5; Length 146407;  
Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACGTTAAGATGAGTC 18  
Db 7456 GACACGTTAAGATGAGTC 7439

RESULT 14

CHKOVALSA/c CHKOVALSA

LOCUS CHKOVALSA

DEFINITION Chicken ovalbumin gene, 5' flank.

ACCESSION M29020

VERSION M29020.1

KEYWORDS Ovalbumin.

SOURCE Gallus gallus (chicken)

ORGANISM Butarvota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 5322)

AUTHORS Kaye,J.S.

JOURNAL Unpublished (1989)

COMMENT Original source text: Chicken DNA, clones PBR ECO RI[6,7]. Draft entry and computer-readable sequence for [1] kindly submitted by J.S.Kaye, 14-Oct-1989.

FEATURES Source Location/Qualifiers

1. . 5322 /organism="Gallus gallus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9031"

ORIGIN

Query Match 80.9%; Score 17.8; DB 5; Length 5322;  
Best Local Similarity 90.5%; Pred. No. 1.6e+02; Mismatches 2; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACACGTTAAGATGAGTC 21  
Db 1539 GAGACGTTAAGATGAGTCAGCA 1519

RESULT 15  
 ABO11563 AE011563 leptospira interrogans serovar lai str. 56601 linear BCT 28-APR-2003  
 LOCUS DEFINITION 372 of 397 of the complete sequence.  
 ACCESSION ABO11563 AE01000  
 VERSION ABO11563.1 GI:24198148  
 KEYWORDS leptospira interrogans serovar lai str. 56601  
 SOURCE leptospira interrogans serovar lai str. 56601  
 ORGANISM Leptospira interrogans serovar lai str. 56601  
 REFERENCE 1 (bases 1 to 10808)  
 AUTHORS Ren, S.X., Fu, G., Jiang, X.G., Zeng, R., Miao, Y.G., Xu, H., Zhang, Y.X.,  
 Xiong, H., Lu, G., Lu, L.F., Jiang, H.Q., Jia, J., Tu, Y.F., Jiang, J.X.,  
 Gu, W.Y., Zhang, Y.Q., Cai, Z., Sheng, H.H., Yin, H.F., Zhang, Y.,  
 Zhu, G.F., Wan, M., Huang, H.L., Qian, Z., Wang, S.Y., Ma, W., Yao, Z.J.,  
 Shen, Y., Qian, B.Q., Xia, Q.C., Guo, X.K., Danchin, A., Saint,  
 Girons, I., Somerville, R.I., Wen, Y.M., Shi, M.H., Chen, Z., Xu, J.G.  
 and Zhao, G.P.  
 TITLE Unique physiological and pathogenic features of Leptospira  
 JOURNAL Nature  
 PUBLISHED 422 (6934), 888-893 (2003)  
 REFERENCE 2 (bases 1 to 10808)  
 AUTHORS Ren, S., Fu, G., Jiang, X., Zeng, R., Xiong, H., Lu, L., Lu, G., Jiang, H.,  
 Ding, Y., Tu, Y., Gu, W., Cai, Z., Sheng, H., Yin, H., Zhang, Y.,  
 Zhao, G., Wang, S., Shen, Y., Qiang, B., Chen, Z., Wen, Y., Xu, J. and  
 Zhao, G.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-MAR-2002) Chinese National Human Genome Center at  
 Shanghai, 250 Bi Bo Road, Shanghai 201203, China  
 COMMENT Updated information will be available at our World Wide Web site  
 (<http://www.chgc.sh.cn/lept/>). Comments to the authors are  
 appreciated.  
 FEATURES source  
 /location/Qualifiers  
 1. 1-10808  
 /organism="Leptospira interrogans serovar lai str. 56601"  
 /mol\_type="genomic DNA"  
 /strain="56601"  
 /serovar="Lai"  
 /db\_xref="taxon:18951"  
 /chromosome="I"  
 /note="Strain 56601 is maintained by the Institute of Epidemiology and Microbiology, Chinese Academy of Preventive Medicine"  
 228..722  
 /locus\_tag="LA4087"  
 228..722  
 /locus\_tag="LA4087"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="AAN51285.1"  
 /db\_xref="GI:24198149"  
 /translation="MFLFWRLRKMKKKILLISITPRSSCKPDPVTEQRKDLQQILMBNNTIKELMKERKEKIPNQIJAHVWLVSLANGGLQSABEMQNSIKDKKEDQDV  
 EKFQAYSSFSSENLAESKLKLAGGTGVNRFPQPMNKWTWSQGTRKIRNPYAPENRGCGDUVH"  
 gene  
 CDS  
 /locus\_tag="LA4088"  
 701..1690  
 /locus\_tag="LA4088"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="AAN51286.1"  
 /db\_xref="GI:24198150"  
 /translation="MWPGSTNVECLDEKIKTKKGIGSLGRCPANCQNLQHN  
 SEENSCIGEKGKMAFEGKAGINGASAFQALQPPVIEFISKGRVYLETILYFLS  
 IIEPKVKEVIDSFFNKLKLKLYTNTYQFOTDDEBKLSKQAFDLDRENAYWGL  
 SPERCSFTAYLKAKDPRFASPLTVPSEAVSDIRNLAGLAVNEEKEKVLISKDG  
 YELLIQIPIGYRHLRSDFPEPFLLSTMEFLVLRQQLTESHAILEKYGKGL  
 NHQSISFGDSLVSLELAISMELGIEBEKEILGRSNKLKELLKHKHLKNEIT"  
 gene  
 CDS  
 /locus\_tag="LA4093"  
 6247..6918  
 /locus\_tag="LA4093"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="AAN51281.1"  
 /db\_xref="GI:24198155"  
 /locus\_tag="LA4093"  
 6247..6918  
 /locus\_tag="LA4093"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="AAN51281.1"

gene  
 CDS  
 /complement(1699..2523)  
 /locus\_tag="LA4089"  
 /locus\_start=1  
 /product="phenazine biosynthesis protein, PhzF family"  
 /protein\_id="AAN51287.1"  
 /translation="MSILQNECSKAIFMLQIIDSFAEKVFOGNPAWCPIQEWI  
 SDWMQKIASNLSEIAFFLEGSSYRIRPTPAREVDIGCHATLAANHFLFEGEV  
 IGDKLKQFQSLGEBLNFKKENMLYDPPSKRPEKETPEKILNSPSISPKVFNSRDY  
 MILYENENOLKLEYSVEGVTEBLNTGIIATCNGDSYDLSRFAPSGLYEFVTG  
 SAHCTLIPWBRSLGKRMKAYQTSARGGKLFCDLGEVRIGGTCVPUYLDGWBI"  
 2752..2863  
 /locus\_tag="LA4090"  
 2752..2863  
 /locus\_tag="LA4090"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="AAN51288.1"  
 /db\_xref="GI:24198152"  
 /translation="MKKFFFVPTLEFVHICPGMSAALPQKPLFRNQFILL"  
 /complement(3535..5007)  
 /genes="pur"  
 /locus\_tag="LA4092"  
 /complement(3535..5007)  
 /genes="pur"  
 /locus\_tag="LA4092"  
 /EC\_number="2.4.2.14"  
 /note="Glutamine phosphoribosylpyrophosphate amidotransferase"  
 /codon\_start=1  
 /transl\_table=11  
 /product="amidophosphoribosyltransferase"  
 /protein\_id="AAN51289.1"  
 /db\_xref="GI:24198153"  
 /translation="MSRIPDKSILRROAQDKRKECAIFGFNFNSBASNFTYGLYS  
 MHQGQESSGIVSSDGEHLYRAGMGLVHTIETTKRKELOQNAATCHNRYTGTGAF  
 LRNAQPLRVSHIGAVASLHNNGNWLNSWLRROLEKGSIFOTUTSEVYLWMAHG  
 ETDTLSAASSAALKVGRGSSLVLTKVNDRRGRDNGRRPDRPDRPDRPDRPDRP  
 ADITDTKVRDVEPGEMVWDQGMNSYYPKPKASLICIFETVIFARPSSISIGS  
 VYKRKNIGRFKARLPPADWVTPVDSANTAAALGTAEEGGSISYOSGLISHYHT  
 FIEDQOKURDFGAKYKNUVPPDQKTRVYKTRVYKTRVYKTRVYKTRVYKTRVY  
 IYVSAAPPTISPCYGDIDPTHELNLIAHITIEBIRKLRVSIAYLVSVESMMRAWDH  
 KGGGFCNACFTACOPVFEPOSELGNQNSLKEYQVEERVVV"  
 4950..6215  
 /locus\_tag="LA4091"  
 4950..6215  
 /locus\_tag="LA4091"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Rabonuclelease D"  
 /protein\_id="AAN51290.1"  
 /db\_xref="GI:24198154"  
 /translation="MWHPGIVLFLWYREFDSFAFLSLKSQVAFMOINSYIVWD  
 TARSQIOLVINTSISQDSDISVDESESSGYTFESKCVIQISLKGKQVYDPLQLQES  
 LGENLEDFDKLKLKPSAIDTAKLKDGFGRQNTADTGSSRJUDHEQSLTIVY  
 YPKIKLKKKEQGSNWEKRLQVQALDQVYLTWENKEKSLIKRNLKLYERAISE  
 FRIKASSEPGSGNSISMDKPFILLYSABERFRYLTIVFDDESRKUNKAPPRVN  
 FERKFRKIKWRTTMSIRRMQHOMMSKNTMELLSQKNPNTDLELRKMLNLFSEWKVH  
 6247..6918  
 /locus\_tag="LA4093"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="AAN51281.1"

```



complement (8151..-8264)



/locus_tag="LA4094" (8151..8264)



/locus_tag="LA4094"



CDS /codon_start=1



/transl_table=11



/product="hypothetical protein"



/protein_id="RAN5292_1"



/db_xref="GI:24198156"



/translation="MGGVVKLDSLTKTYLCKITNWLKETKFVGTHINVLF"



complement (8473..-8577)



/locus_tag="LA4095" (8473..-8577)



complement (8473..-8577)



/locus_tag="LA4095"



CDS /codon_start=1



/transl_table=11



/product="hypothetical protein"



/protein_id="RAN5293_1"



/db_xref="GI:24198157"



/translation="MTIHNFLCVSSYIIGPSDOFFENDSKULLFHLT"



8647..-8820



/locus_tag="LA4096"



8647..-8820



/locus_tag="LA4096"



CDS /codon_start=1



/transl_table=11



/product="hypothetical protein"



/protein_id="RAN5294_1"



/db_xref="GI:24198158"



/translation="MKIIKLNFLQKEIRSSFLKESFFQBBEWYQSLVNRPGIEERIPYFK
TKAENKKAAVVR"



8817..-9212



/locus_tag="LA4097"



8817..-9212



/locus_tag="LA4097"



CDS /codon_start=1



/transl_table=11



/product="conserved hypothetical protein"



/protein_id="RAN5295_1"



/db_xref="GI:24198159"



/translation="MNNEBKAKLLETLIDLKHIQSQSSNSGSDYKVULVNPIFY
FGSALLFFVYWWRRQRRIKIKADYKKETFDLYTYSFFLGLLUTFVGIALSIGFISV
LGGSAMUGGLVPLGTGIGLCCYKFSSG"



9212..-9790



/gene="tpeB2"



/locus_tag="LA4098"



9212..-9790



/gene="tpeB2"



/locus_tag="LA4098"



/codon_start=1



/transl_table=11



---



Query Match Score 17.8; DB 1; Length 10808;  

  Best Local Similarity 90.5%; Pred. No. 1.6e+02;  

  Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



|    |      |                      |      |
|----|------|----------------------|------|
| QY | 1    | GACACCTAGATGAGTCATCA | 21   |
| Db | 5999 | GATCAGAAGATGAGTCATCA | 6019 |



Search completed: March 1, 2005, 00:29:25  

  Job time : 648.714 secs


```

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: February 28, 2005, 23:35:01 ; Search time 2081.3 Seconds  
(without alignments)  
548.661 Million cell updates/sec

Title: US-10-086-062-4  
Perfect score: 30  
Sequence: 1 ctggacccttcgactcgagatccgt 30

Scoring table: IDENTITY\_NUC  
Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 68479088  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : EST.\*  
1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_htc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gbs1:\*

9: gb\_gbs2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

#### ALIGNMENTS

RESULT 1	BW905373	AGENCOURT	BW905373	6698805 NIH_MGC_72 Homo sapiens	1091 bp mRNA clone IMAGE:5558278	EST	12-MAR-2002
LOCUS				5' mRNA sequence.			
DEFINITION				BW905373			
ACCESSION				6698805 NIH_MGC_72			
VERSION				1			
KEYWORDS				GI:19355752			
SOURCE				EST.			
ORGANISM				Homo sapiens (human)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE				1 (bases 1 to 1091)			
AUTHORS				NIH-MGC <a href="http://mgc.ncbi.nih.gov/">http://mgc.ncbi.nih.gov/</a>			
TITLE				National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT				Unpublished (1999)			
Contact: Robert Straubberg, Ph.D. Email: cgs@nih.gov							
Tissue Procurement: ARCC/DCDP/NP							
CDNA Library Preparation: Life Technologies, Inc.							
CDNA Sequencing by: AgenCourt Bioscience Corporation							
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:							
http://image.llnl.gov							
Plate: LLM1M2280 row: P column: 23							
High quality sequence stop: 573.							
FEATURES				Location/Qualifiers			
SOURCE				1. .1091			
				/organism="Homo sapiens"			
				/mol_type="mRNA"			
				/db_xref="taxon:9606"			
				/clone=IMAGE:5558278"			
				/issue_type="melanotic melanoma"			
				/lab_host="DH10B (phage-resistant)"			
				/clone_lib="NIH_MGC_72"			
				/clone_lib="NIH_MGC_72"			
				/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."			
C	18	19.4	64.7	469	4 BI516790 BB160023B		
C	19	19.4	64.7	511	4 BI514171 BB160014B		
C	20	19.4	64.7	541	4 BI513884 BB160014A		
C	21	19.4	64.7	564	4 BI513971 BB160014A		
C	22	19.4	64.7	598	1 A193892 701515328		
C	23	19.4	64.7	688	2 AW47649 hb06603.X		
C	24	19.4	64.7	893	7 CO918811 AGENCOURT		
C	19	63.3	390	5 BX087393			
C	19	63.3	394	5 BX087393			
C	20	63.3	453	5 BX865652			
C	21	63.3	457	6 CB739770 AMGNNC: N			
C	22	63.3	467	8 A2651293 IM0521J23			
C	23	63.3	498	5 BX301324			

ORIGIN	Query Match	68.7%	Score	20.6	DB	5	Length	1091;
	Best Local Matches	85.2%	Pred. No.	3.1e+02;				
	23;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps 0;
QY	1	CTGGACCCCTCGACTCGAGAGTTCC	27					

**Db** 900 |||||||CTGGACCCCTCTCCACTCGGAATCCC 926

**RESULT 2**

**LOCUS** BG430455/c 804 bp mRNA linear EST 14-MAR-2001

**DEFINITION** 602502021FL NIH\_MGC\_75 Homo sapiens cDNA clone IMAGE:4615756 5', mRNA sequence.

**REFERENCE** 1 (bases 1 to 804)

**AUTHORS** NIH-MGC http://mgc.ncbi.nih.gov/

**ACCESSION** BG430455

**VERSION** BG430455.1 GI:13337063

**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**COMMENT** NIH-MGC http://mgc.ncbi.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Daniel S. Gerhard, Ph.D.

**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)

**JOURNAL** Unpublished (1999)

**COMMENT** Contact: Robert Strausberg, Ph.D.

**Email:** cgsbps-r@mail.nih.gov

**Tissue** Procurement: Clontech Laboratories, Inc.

**Procurement** DNA Library Preparation: Clontech Laboratories, Inc.

**Procurement** DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU)

**Clon** distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNU at: http://image.llnl.gov/Plate: LNUM1169 row: i column: 05

**High quality sequence stop:** 742.

**FEATURES**

**source**

1. . 804

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4615756"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NIH\_MGC\_75"

/note="Organ: kidney; Vector: PDNR-LIB (Clontech); Site 1: SfiI (ggccgcctccggc); Site 2: SfiI (ggccatttagggcc); 5', and 3' adaptors were used in cloning as follows: 5', adaptor sequence: 5'-ATTCAGAGCCGAGCGAGCATG-3' (where B = A, C, or G and N = A, C, G or T), Average insert size 1.65 kb (range 0.5-4.0 kb), 15/15 colonies contained insert B by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC library."

**ORIGIN**

Query Match Best Local Similarity 68.0%; Score 20.4; DB 4; Length 804; Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

**QY** 1 CTGACCCCTCTGACTCGAGAGTCCTACT 30

**Db** 559 CTGAGCCCTCTCGCTGGAGACTTCCCT 530

**RESULT 3**

**LOCUS** CK873283/c 877 bp mRNA linear EST 09-MAR-2004

**DEFINITION** AGENCOURT\_19436603 NIH\_ZGC\_16 Danio rerio cDNA clone IMAGE:7215126

**ACCESSION** CK873283

**VERSION** CK873283.1 GI:45289381

**KEYWORDS** EST.

**SOURCE** Danio rerio (zebratfish)

**ORGANISM** Danio rerio; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

**REFERENCE** 1 (bases 1 to 877)

**AUTHORS** NIH-MGC http://mgc.ncbi.nih.gov/

**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)

**JOURNAL** Unpublished (1999) Contact: Daniel S. Gerhard, Ph.D.

**COMMENT** Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgsbps-r@mail.nih.gov

**Tissue** Procurement: Will Talbot

**Procurement** CDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano

**Procurement** CDNA Sequencing by: The I.M.A.G.E. Consortium (LNU)

**Clon** distribution: Agencourt Bioscience Corporation found through the I.M.A.G.E. Consortium/LNU at: http://image.llnl.gov/Plate: LNUM1511 row: h column: 04

**High quality sequence start:** 13

**High quality sequence stop:** 685.

**FEATURES**

**source**

1. . 877

/organism="Danio rerio"

/mol\_type="mRNA"

/db\_xref="taxon:7953"

/clone="IMAGE:7215126"

/tissue\_type="13 pooled, includes stomach, intestine, liver and pancreas"

/lab\_host="DH10B Tona"

/clone\_lib="NIH\_ZGC\_16"

/note="Organ: gut; Vector: pME18-FL3; Site 1: DraIII; Site 2: DraIII; 1st strand cDNA was primed with an oligo(dT) primer [GGGCTGAAAGACGCCATATGGCCTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [GGCCUAGUGG], digested and directionally cloned into distinct DAI1 sites of the pME18-FL3. Library was selected for 1.0 kb, with a average insert size of ~1.2kb. Library constructed by Yutaka Suzuki (University of Tokyo Institute of Medical Science). Custom primers recommended for sequencing: 5' end primer 5'-GGAGTGCGCTTACTGTA-3' and 3' end primer 5'-CGAACCTGAGCTCGAGACA-3'. Note: This is a zebrafish Gene Collection (ZGC) library."

**ORIGIN**

Query Match Best Local Similarity 67.3%; Score 20.2; DB 7; Length 877; Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

**QY** 2 TGGACCCCTCTGACTCGAGAGTC 26

**Db** 817 TGGACCTGAGCTCGAGAGTC 793

**RESULT 4**

**LOCUS** BG474756 1218 bp mRNA linear EST 21-MAR-2001

**DEFINITION** 602517583FL NIH\_MGC\_16 Homo sapiens cDNA clone IMAGE:4649109 5', mRNA sequence.

**ACCESSION** BG474756

**VERSION** BG474756.1 GI:13407070

**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**COMMENT** NIH-MGC http://mgc.ncbi.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D.

**Email:** cgsbps-r@mail.nih.gov

**Tissue** Procurement: ATCC

**Procurement** CDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at: image.lnl.gov  
 Plate: ILCM128 row: f column: 22

## FEATURES source

High quality sequence Stop: 1.  
 Location/Qualifiers

1..1218

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_id="IMAGE:4649109"

/tissue\_type="retinoblastoma"

/lab\_host="DHIOB (phage-resistant)"

/note="Organ: eye; Vector: pORB7; Site\_1: XbaI; Site\_2:

cloned into EcoRI/XbaI sites using the following 5'

adaptor: GGCACCG (G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC library."

## ORIGIN

Query Match Score 67.3%; DB 4; Length 1218;  
 Best Local Similarity 88.0%; Pred. No. 4.6e+02;  
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TGGACCCCTCTGACTGGAGAGTC 26  
 Db 769 TGGAACCCCTCTGACTGGAGAGTC 793

## RESULT 5

LOCUS CN127674/C 838 bp mRNA Linear EST 01-APR-2004  
 DEFINITION RHOH1\_24 A01\_93 A002 Acid- and alkaline-treated roots Sorghum  
 bicolour cDNA clone RHOH1\_24\_A01\_A002 5', mRNA sequence.

## ACCESSION CN127674

## VERSION CN127674.1

## KEYWORDS EST

## SOURCE Sorghum bicolor (sorghum)

## ORGANISM Sorghum bicolor

## REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; Liliopsida; Poales; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

AUTHORS Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C., Sun,F., Sullivan,R., Lim,S., Eastman,A. and Pratt,L.H.  
 TITLE An EST database from Sorghum: acid- and alkaline-treated roots  
 JOURNAL Unpublished (2003)  
 COMMENT Other ESTs: RHOH1\_24 A01\_b1 A002  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mpratt@uga.edu

Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
 the Human Genome Center, University of Tokyo Institute of Medical  
 Science; plant material and RNA prepared at Texas A & M University;  
 sequencing done in the Laboratory for Genomics and Bioinformatics,  
 University of Georgia. Sequence ends have been trimmed to exclude  
 vector and regions below Phred quality 16. Three-prime sequences  
 are presented as their reverse complement and have been trimmed to  
 exclude polyA.  
 Seq primer: S95 (CTCTCTCTAAAGCTGCG).  
 Location/Qualifiers

1..138

/organism="Sorghum bicolor"

## FEATURES source

Source

## ORIGIN

Query Match Score 66.7%; DB 7; Length 838;  
 Best Local Similarity 82.1%; Pred. No. 5.5e+02;  
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTGACTGGAGAGTC 28  
 Db 526 CTGGACCCACAAGCATCGAGAGTC 499

## RESULT 6

LOCUS CR429361/c 634 bp DNA linear Canis familiaris genomic,  
 DEFINITION tigr-gba-dog-1700362934386 Dog library Canis familiaris genomic,  
 genomic survey sequence.

## ACCESSION CR429361

## VERSION CE429361.1

## KEYWORDS GSS

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;

Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

## REFERENCE 1 (bases 1 to 634)

Kirkness,E.F., Barna,V., Halpern,A.L., Levy,S., Remington,K.,

Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Frazer,C.M. and

Venter,J.C.

The dog genome: survey sequencing and comparative analysis

Science 301 (5641), 1898-1903 (2003)

## TITLE 2875432

## JOURNAL 14512627

COMMENT Contact: Kirkness EF

The Institute for Genomic Research, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirkness@tigr.org

Class: shotgun.

Location/Qualifiers

1..634

/organism="Canis familiaris"

/mol\_type="genomic DNA"

/strain="Standard Poodle"

/db\_xref="taxon:9615"

/clone\_id="dog Library"

/note="Site\_1: BstXI; libraries were prepared from

peripheral Blood"

## ORIGIN

Query Match Score 65.3%; DB 9; Length 634;  
 Best Local Similarity 84.6%; Pred. No. 8.2e+02;  
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TGGACCCCTCTGACTGGAGAGTC 27  
 Db 368 TGGACCCCTCTGGCTAGAGAGTC 343

RESULT 7		SOURCE	Leishmania braziliensis
BFB89056/C	BB889056	ORGANISM	Leishmania braziliensis
LOCUS	MN4-TN0116-271100-202-e05	DEFINITION	Bukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania; Leishmania braziliensis species complex;
ACCESSION	BB889056	REFERENCE	1 Laurentino,E.C., Ruiz,J.C. and Cruz,A.K.
VERSION	BB889056.1	AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., de Oliveira,P.S., Bucher,P., Jongenel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
KEYWORDS		TITLE	GSS analysis of the Leishmania braziliensis genome
SOURCE	Homo sapiens (human)	JOURNAL	Unpublished
ORGANISM		REFERENCE	2 (bases 1 to 168)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Homidae; Homo.		AUTHORS	Cruz,A.K.
PUBLISHED	10/3/2000	TITLE	Direct Submission
COMMENT		JOURNAL	Submitted (17-TUN-2003) Cruz A.K., University of Sao Paulo, Department of Molecular and Cell Biology, FMERP, Avenida Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., de Oliveira,P.S., Bucher,P., Jongenel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.	COMMENT	Clone requests: akcruz@fmpur.usp.br.
TITLE		FEATURES	Location/Qualifiers
JOURNAL		source	1..168
MEDLINE			/organism="Leishmania_braziliensis"
PUBMED	10373800		/mol_type="genomic DNA"
COMMENT			/strain="MOM/BR/75/M2904"
AUTHORS	Contact: Simpson A.J.G.		/db_xref="ttaxon:5660"
Laboratory of Cancer Genetics			/clone="LBAF041G06"
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil			
TELEPHONE	+55-11-2704922		
FAX	+55-11-2707001		
EMAIL	asimpson@ludwig.org.br		
FEATURES	This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL ( <a href="http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MN4&amp;t2=MN4-TN0116-Seq_primer_Puc_18_forward">http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MN4&amp;t2=MN4-TN0116-Seq_primer_Puc_18_forward</a> )		
source	High quality sequence start: 68		
LOCATION/QUALIFIERS	Location/Qualifiers		
1..123			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:3606"			
/dev_stage="adult"			
/clone_lib="TN0116"			
/note="Organ: testis normal; Vector: puc18; Site 1: Small; Site 2: Small. A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the PUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."			
ORIGIN			
Query Match	64.7%	Score	19.4;
Best Local Similarity	79.3%	DB	9;
Matches	23;	Length	168;
Conservative	0;	Pred. No.	9.6e+02;
Mismatches	6;	Mismatches	6;
Indels	0;	Indels	0;
Gaps	0;	Gaps	0;
REFERENCE	BI055318/C	DEFINITION	RCO-GN0233-060201-031-f07 GN0233 Homo sapiens cDNA, mRNA sequence.
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., de Oliveira,P.S., Bucher,P., Jongenel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.	VERSION	BI055318.1
KEYWORDS		FEATURES	EST.
SOURCE	Homo sapiens (human)	ORGANISM	
ORGANISM	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Homidae; Homo.	REFERENCE	1 Laurentino,E.C., Ruiz,J.C. and Cruz,A.K.
PUBLISHED	10/3/2000	COMMENT	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., de Oliveira,P.S., Bucher,P., Jongenel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
COMMENT	Contact: Simpson A.J.G.	FEATURES	Laboratory of Cancer Genetics
AUTHORS	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil	REFERENCE	2 (bases 1 to 168)
TELEPHONE	+55-11-2704922	COMMENT	Submitted (17-TUN-2003) Cruz A.K., University of Sao Paulo, Department of Molecular and Cell Biology, FMERP, Avenida Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL
FAX	+55-11-2707001	FEATURES	Clone requests: akcruz@fmpur.usp.br.
EMAIL	asimpson@ludwig.org.br	REFERENCE	Cruz,A.K.
FEATURES	This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL ( <a href="http://www.ludwig.org.br/scripts/gethtml2.pl?ti=RC0-GN0233-060201-031-f07&amp;t2=02-02-06&amp;t4=1">http://www.ludwig.org.br/scripts/gethtml2.pl?ti=RC0-GN0233-060201-031-f07&amp;t2=02-02-06&amp;t4=1</a> )	COMMENT	
source	High quality sequence start: 69	FEATURES	
LOCATION/QUALIFIERS	Location/Qualifiers	source	
1..240			

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dv_stage="adult"
/clone_lib="GN0233"
/note="Organ: placenta-normal; Vector: pUC18; Site_1: small; Site_2: Small; A mini-library was made by cloning products derived from OREGES PCR (U.S. Letters Patent application No. 1,967,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN					
Query	CTGGACCCCTCTGAGCTGGAGTTCCGC	64.7%	Score 19.4;	DB 4;	Length 240;
Best Local Similarity	79.3%	Pred. No. 9.7e+02;			
Matches	23;	Conservative	0;	Mismatches	6;
Ov		29			
Ds	CTGGAGCTGTGAGCTTAGAGATCCTC	55			
		27			
			Indels	0;	Gaps
				0;	

**RESULT 11**  
 BFR822389/c  
**LOCUS** BFR822389  
**DEFINITION** M1-RT0078-061200-002-a10 RT0078 Homo sapiens cDNA, mRNA sequence.  
**ACCESSION** BFR822389  
**VERSION** BFR822389.1 GI:12162356  
**KEYWORDS** EST  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens

RESULT 10  
 BF85649/C  
 LOCUS  
 DEFINITION BF856649 245 bp mRNA linear EST 16-JAN-2001  
 ACCESSION MR1\_FT0200-311000-001-a04  
 VERSION BF856649  
 KEYWORDS GI:12244393  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Homo sapiens

**REFERENCE**  
**AUTHORS**  
 Butarweka; Metazoa; Chordata; Craniata; Vertebrata; Bivalvia;  
 Mammalia; Eutheria; Primates; Veribranchii; Hominidae; Homo.  
 1 (bases 1 to 254)  
 Dias Nevo, E., Garcia Correa, R., Verjovoski-Almeida, S., Briones, M.R.,  
 Niagai, M.A., da Silva, W. Jr., Zogo, M.A., Bordin, S., Costa, F.P.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
**TITLE**  
 shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

**REFERENCE**  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**AUTHORS**  
 1 (bases 1 to 245)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordim,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
 Brunstein,A., de Oliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
**TITLE**  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
**JOURNAL**  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
**MEDLINE**  
 20202663  
**PUBMED**  
 10737800  
**COMMENT**  
 Contact: Simpson A.J.G.

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 2029663  
PUBMED 10737800  
COMMENT  
Contact: Simpson A.J.G.  
Ludwig Institute for Cancer Research  
Ludwig Institute for Cancer Genetics<sup>B</sup>  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, São Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR1&t2=MR1-RT0078-061200-002-a10&t3=2000-12-06&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 40  
High quality sequence stop: 235.

Email: abimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR1&t2=MR1-FT02000-311001-a04&t3=2000-10-31&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 34  
High quality sequence stop: 157.  
location/Qualifira

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606".
/dev_stage="adult"
/clone_lib="RT0078"
/note="Organ: kidney tumor; Vector: pUC18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

/clone lib="PT0200"  
/note="Organ: prostate tumor; Vector: puc18; Site\_1: Smal;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the PUC 18 vector. Reverse transcription of

Db 29 CTGGAGCTGCTGACTCTAGAGATCCAC 1

RESULT 12

B1516790

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

APIS MELLIFERA (HONEY BEE)

APIS MELLIFERA

NEOPTERA; METACOEA; ARTHROPODA; HEXAPODA; INSECTA; PTERYGOTA;

NEOPTERA; ENDOPTERYGOTA; HYMENOPTERA; APOCITA; ACULEATA; APOIDEA;

APODIAE; APIS.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: GENE E. ROBINSON

DEPARTMENT OF ENTOMOLOGY

UNIVERSITY OF ILLINOIS

505 S. GOODWIN AVE., URBANA, IL 61801, USA

TEL: 217 265 0309

FAX: 217 244 3499

EMAIL: genrobb@life.uiuc.edu

THIS RESEARCH WAS FUNDED BY THE UNIVERSITY OF ILLINOIS CRITICAL

RESEARCH INITIATIVES FUND AND A BURROUGHS-WELLCOME TRUST INNOVATION

AWARD IN FUNCTIONAL GENOMICS TO G.E. ROBINSON AND AN NSF

POSTDOCTORAL FELLOWSHIP IN BIOINFORMATICS TO C.W. WHITFIELD.

PCR PRIMERS

FORWARD: TAAATGGACTCACTATAGG

BACKWARD: ATTAAACCCTGACTTAAG

PLATE: BB160023B10 ROW: C COLUMN: 10

SEQ PRIMER: AGCGGTAACTAACATCACAGGA

HIGH QUALITY SEQUENCE STOP: 469.

FEATURES

SOURCE

APIS MELLIFERA (HONEY BEE)

NEOPTERA; ENDOPTERYGOTA; HYMENOPTERA; APOCITA; ACULEATA; APOIDEA;

APODIAE; APIS.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: GENE E. ROBINSON

DEPARTMENT OF ENTOMOLOGY

UNIVERSITY OF ILLINOIS

505 S. GOODWIN AVE., URBANA, IL 61801, USA

TEL: 217 265 0309

FAX: 217 244 3499

EMAIL: genrobb@life.uiuc.edu

THIS RESEARCH WAS FUNDED BY THE UNIVERSITY OF ILLINOIS CRITICAL

RESEARCH INITIATIVES FUND AND A BURROUGHS-WELLCOME TRUST INNOVATION

AWARD IN FUNCTIONAL GENOMICS TO G.E. ROBINSON AND AN NSF

POSTDOCTORAL FELLOWSHIP IN BIOINFORMATICS TO C.W. WHITFIELD.

PCR PRIMERS

FORWARD: TAAATGGACTCACTATAGG

PLATE: BB160023B10 ROW: C COLUMN: 10

SEQ PRIMER: AGCGGTAACTAACATCACAGGA

HIGH QUALITY SEQUENCE STOP: 511.

FEATURES

SOURCE

APIS MELLIFERA (HONEY BEE)

NEOPTERA; ENDOPTERYGOTA; HYMENOPTERA; APOCITA; ACULEATA; APOIDEA;

APODIAE; APIS.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: GENE E. ROBINSON

DEPARTMENT OF ENTOMOLOGY

UNIVERSITY OF ILLINOIS

505 S. GOODWIN AVE., URBANA, IL 61801, USA

TEL: 217 265 0309

FAX: 217 244 3499

EMAIL: genrobb@life.uiuc.edu

THIS RESEARCH WAS FUNDED BY THE UNIVERSITY OF ILLINOIS CRITICAL

RESEARCH INITIATIVES FUND AND A BURROUGHS-WELLCOME TRUST INNOVATION

AWARD IN FUNCTIONAL GENOMICS TO G.E. ROBINSON AND AN NSF

POSTDOCTORAL FELLOWSHIP IN BIOINFORMATICS TO C.W. WHITFIELD.

PCR PRIMERS

FORWARD: TAAATGGACTCACTATAGG

PLATE: BB160023B10 ROW: C COLUMN: 10

SEQ PRIMER: AGCGGTAACTAACATCACAGGA

HIGH QUALITY SEQUENCE STOP: 511.

FEATURES

SOURCE

APIS MELLIFERA (HONEY BEE)

NEOPTERA; ENDOPTERYGOTA; HYMENOPTERA; APOCITA; ACULEATA; APOIDEA;

APODIAE; APIS.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: GENE E. ROBINSON

DEPARTMENT OF ENTOMOLOGY

UNIVERSITY OF ILLINOIS

505 S. GOODWIN AVE., URBANA, IL 61801, USA

TEL: 217 265 0309

FAX: 217 244 3499

EMAIL: genrobb@life.uiuc.edu

THIS RESEARCH WAS FUNDED BY THE UNIVERSITY OF ILLINOIS CRITICAL

RESEARCH INITIATIVES FUND AND A BURROUGHS-WELLCOME TRUST INNOVATION

AWARD IN FUNCTIONAL GENOMICS TO G.E. ROBINSON AND AN NSF

POSTDOCTORAL FELLOWSHIP IN BIOINFORMATICS TO C.W. WHITFIELD.

PCR PRIMERS

FORWARD: TAAATGGACTCACTATAGG

PLATE: BB160023B10 ROW: C COLUMN: 10

SEQ PRIMER: AGCGGTAACTAACATCACAGGA

HIGH QUALITY SEQUENCE STOP: 511.

RESULT 14

B1513884

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

APIS MELLIFERA (HONEY BEE)

APIS MELLIFERA

NEOPTERA; METACOEA; ARTHROPODA; HEXAPODA; INSECTA; PTERYGOTA;

NEOPTERA; ENDOPTERYGOTA; HYMENOPTERA; APOCITA; ACULEATA; APOIDEA;

APODIAE; APIS.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: GENE E. ROBINSON

DEPARTMENT OF ENTOMOLOGY

UNIVERSITY OF ILLINOIS

505 S. GOODWIN AVE., URBANA, IL 61801, USA

TEL: 217 265 0309

FAX: 217 244 3499

EMAIL: genrobb@life.uiuc.edu

THIS RESEARCH WAS FUNDED BY THE UNIVERSITY OF ILLINOIS CRITICAL

RESEARCH INITIATIVES FUND AND A BURROUGHS-WELLCOME TRUST INNOVATION

AWARD IN FUNCTIONAL GENOMICS TO G.E. ROBINSON AND AN NSF

POSTDOCTORAL FELLOWSHIP IN BIOINFORMATICS TO C.W. WHITFIELD.

PCR PRIMERS

FORWARD: TAAATGGACTCACTATAGG

PLATE: BB160023B10 ROW: C COLUMN: 10

SEQ PRIMER: AGCGGTAACTAACATCACAGGA

HIGH QUALITY SEQUENCE STOP: 511.

Db

ACCESSION	BIS13884	ORGANISM	Apis mellifera
VERSION	BIS13884.1	EST:	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis
KEYWORDS		AUTHORS	Apidae; Apis
SOURCE	Apis mellifera	REFERENCE	1 (bases 1 to 564)
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis	JOURNAL	Whittfield, C.W., Band, M.R., Ronaldo, M.F., Kumar, C.G., Liu, L., Pardinas, J.J., Robertson, H.M., Soares, B. and Robinson, G.E.
REFERENCE	1 (bases 1 to 541)	MEDLINE	Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee
AUTHORS	Whittfield, C.W., Band, M.R., Ronaldo, M.F., Kumar, C.G., Liu, L., Pardinas, J.J., Robertson, H.M., Soares, B. and Robinson, G.E.	JOURNAL	Genome Res. 12 (4), 555-566 (2002)
TITLE	Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee	MEDLINE	11932240
COMMENT	Contact: Gene E. Robinson Department of Entomology University of Illinois 505 S. Goodwin Ave., Urbana, IL 61801, USA Tel: 217 265 0309 Fax: 217 244 3499 Email: generobi@life.uiuc.edu	COMMENT	Contact: Gene E. Robinson Department of Entomology University of Illinois 505 S. Goodwin Ave., Urbana, IL 61801, USA Tel: 217 265 0309 Fax: 217 244 3499 Email: generobi@life.uiuc.edu
JOURNAL	Genome Res. 12 (4), 555-566 (2002)	RESEARCHERS	This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation Award in Functional Genomics to G.E. Robinson and an NSF Postdoctoral Fellowship in Bioinformatics to C.W. Whittfield.
MEDLINE	21929762	PRIMER	FORWARD: TAATAGCAGTCACTATAGG BACKWARD: ATTAACCCCTCACTAAG
PUBMED	11932240	PLATE	BB160014A10 row: F column: 07
KEYWORDS	Seq primer: ACGGGATAACATTACACAGGA	SEQ PRIMERS	FORWARD: TAATAGCAGTCACTATAGG BACKWARD: ATTAACCCCTCACTAAG
SOURCE	High quality sequence stop: 541. Location/Qualifiers	PLATE	BB160014A11 row: F column: 07
FEATURES	1. . 541  /organism="Apis mellifera" /mol_type="mRNA" /strain="mixed" strains of European bees, predominantly A.m. ligustica" /db_xref="taxon:7460" /clone="BB160014A10F07" /sex="female" /tissue_type="brain" /dev_stage="adult worker honey bee" /lab_host="DH10B" /clone lib="Bee Brain Normalized Library, BB16" /note="Organ: brain; Vector: pRTT3-Pac; Site: 1: EcoRI; Site: 2: NotI; The BB16 library was contributed by the Soares laboratory and it was constructed and normalized as described by Ronaldo, M.F., Lennon, G. and Soares, M.B. (1996). Genome Research 6(9): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups." ORIGIN Query Match 64.7%; Score 19.4; DB 4; Length 541; Best Local Similarity 95.2%; Pred. No. 9.9e+02; Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0; OY 7 CCCTCTGACTCGAGAGTTCC 27 Db 182 CCCTCTGACTCGAGAGTTCC 202	FEATURES	1. . 564  /organism="Apis mellifera" /mol_type="mRNA" /strain="mixed" strains of European bees, predominantly A.m. ligustica" /db_xref="taxon:7460" /clone "BB160014A1F07" /sex="female" /tissue_type="brain" /dev_stage="adult worker honey bee" /lab_host="DH10B" /clone lib="Bee Brain Normalized Library, BB16" /note="Organ: brain; Vector: pRTT3-Pac; Site: 1: EcoRI; Site: 2: NotI; The BB16 library was contributed by the Soares laboratory and it was constructed and normalized as described by Ronaldo, M.F., Lennon, G. and Soares, M.B. (1996). Genome Research 6(9): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups." ORIGIN Query Match 64.7%; Score 19.4; DB 4; Length 564; Best Local Similarity 95.2%; Pred. No. 9.9e+02; Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0; OY 7 CCCTCTGACTCGAGAGTTCC 27 Db 182 CCCTCTGACTCGAGAGTTCC 202
RESULT	15	SEARCH	Search completed: March 1, 2005, 01:58:25 Job time : 2088.63 secs
B1513971	B1513971 564 bp mRNA Linear EST 08-APR-2002	RESULTS	
LOCUS	BB160014A11F07.5 Bee Brain Normalized Library, BB16 Apis mellifera	DEFINITION	
DEFINITION	CDNA clone BB160014A11F07 5, mRNA sequence.	ACCESSION	BIS13971
ACCESSION	BIS13971.1 GT:15364345	VERSION	BIS13971.1
VERSION	EST: Apis mellifera (honey bee)	KEYWORDS	SOURCE

*This Page Blank (uspto)*

OM nucleic - nucleic search, using sw model	Run on:	February 28, 2005, 23:26:41 ; Search time 261.818 Seconds (without alignments) 678.304 Million cell updates/sec
Title:	Perfect score:	US-10-086-062-4
Sequence:	1ctggacacctctgactcgagatgttcgcgt 30	
Scoring table:	IDENTITY_NUC Gapop 10.0 , Gapext 1.0	
Searched:	439206 seqs, 2959870667 residues	
Total number of hits satisfying chosen parameters:	8780412	
Minimum DB seq length:	0	
Maximum DB seq length:	200000000	
Post-processing:	Minimum Match 0% Maximum Match 100% Listing first 45 summaries	
Database :	N_Geneseq_16Dec04;*	
1:	geneseqn1980s;*	
2:	geneseqn1990s;*	
3:	geneseqn2000s;*	
4:	geneseqn2001as;*	
5:	geneseqn2001bs;*	
6:	geneseqn2002as;*	
7:	geneseqn2002bs;*	
8:	geneseqn2003as;*	
9:	geneseqn2003bs;*	
10:	geneseqn2003cb;*	
11:	geneseqn2003db;*	
12:	geneseqn2004as;*	
13:	geneseqn2004bs;*	
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	
SUMMARIES		
Result No.	Score	Query Match Length DB ID Description
1	30	100.0 30 6 AAD24069 Ad24069 Maize eng
C 2	19.2	64.0 1500 5 AAS11024 Aas11024 Vibrio ch
C 3	19	63.3 121124 12 ADQ97107 Adq97107 Mouse can
C 4	18.8	62.7 23107 10 ABT4166 Abt4166 Toxicity
C 5	18.8	62.7 23107 9 ADR02762 Ada02762 Human RUN
C 6	18.8	62.7 23107 10 ADR72500 Abb72500 Human Run
C 7	18.8	62.7 23107 10 ADC85242 Adc85242 Human Run
C 8	18.8	62.7 23107 12 ADM74357 Adm74357 Human car
C 9	18.6	62.0 3210 2 ADZ27624 Adz27624 Plasmid S
C 10	18.6	62.0 3452 2 ADZ27623 Adz27623 Plasmid S
C 11	18.4	61.3 319608 3 AHS1601 Ahs1601 Human chr
C 12	18.4	61.3 319608 5 AHS09301 Ahs09301 Human sch
C 13	18.2	60.7 340498 8 AAL52198 Human sec
C 14	18	60.0 432325 10 ADB74382 Adb74382 Mycobacte
C 15	18	60.0 110000 11 ADM27081_10 Continuation (11 o
C 16	17.8	59.3 165 10 ADD49400 Add49400 Human lun
C 17	17.8	59.3 254 12 ADDG0322 Addg0322 Nicotiana
C 18	17.8	59.3 369 11 Abd11207 Abd11207 Pseudomon
C 19	17.8	59.3 426 10 ADD49385 Human lun
C 20	17.8	59.3 449 10 ADD49343 Human lun
ALIGNMENTS		
RESULT 1		
ID	AD24069	standard; DNA; 30 BP.
XX		
AC	AD24069;	
XX		
DT	09-APR-2002	(first entry)
XX		
DE	Maize engineered Ubi-1 promoter	heat shock element #3.
XX		
KW	Gene expression; maize; ubiquitin promoter; Ubi-1; HSE;	
KW	heat shock element; agronomic gene; ds.	
XX		
OS	Zea mays.	
XX		
PH		Location/Qualifiers
FT	misc_feature	1..15
FT		/*tag= "a", heat shock element"
FT	misc_feature	/note= "5", heat shock element"
FT		16..30
FT		/*tag= "b", heat shock element"
FT		/note= "3", heat shock element"
XX		
PN	W0200194394-A2.	
XX		
PD	13-DEC-2001.	
XX		
PP	08-JUN-2001; 2001WO-US018689.	
XX		
PR	09-JUN-2000; 2000US-00590558.	
XX		
PA	(PROD-) PRODIGENE INC.	
PI	Jilka JM, Hood EE, Howard JA;	
XX		
DR	WPI; 2002-122117/16.	
XX		
PT	New promoter sequences for causing expression of a structural gene	
PT	especially agronomic gene or open reading frame in a plant cell,	
PT	comprises engineered versions of the maize ubiquitin promoter.	
PS	Claim 6; Page 54; 68pp; English.	
XX		

The invention relates to a promoter sequence capable of directing expression of a nucleotide sequence in a plant cell, comprising maize ubiquitin (Ubi-1) promoter sequence with a modification so that it does not include two overlapping heat shock elements (HSE) or it directs expression to increase the endosperm/embryo expression ratio of the protein when compared to the ratio from a wild-type ubiquitin promoter. The modified Ubi-1 promoter comprises a deletion of 3', 5' or both HSEs, replacement of HSEs by a trimer of a two non-overlapping/adjacent HSEs, replacement of HSEs by a seed specific element from the promoter of pea lectin gene *psl*, or insertion of a transcription factor binding site in the HSE region. An expression construct comprising modified Ubi-1 promoter is useful for causing expression of a structural gene (agronomic genes) or open reading frame in a plant cell. The modified Ubi-1 promoter increases expression levels beyond those observed with native ubiquitin promoter. The present sequence is maize engineered Ubi-1 promoter with heat shock elements adjacently placed. Note: The present sequence is also shown in claim 26, page 56 of the specification. However, this sequence has an additional nucleotide at the 3' end.

XX SQ Sequence 30 BP; 4 A; 12 C; 7 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 6; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.0012; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTGACTCGAGTCGCT 30  
Db 1 CTGGACCCCTCTGACTCGAGTCGCT 30

RESULT 2

AAS11024/c

ID AAS11024 standard; DNA; 1500 BP.

XX AC AAS11024;

XX DT 11-SEP-2003 {revised}

DT 24-OCT-2001 {first entry}

DE Vibrio cholera 16S ribosomal RNA gene.

KW Antisense; bacterial 16S ribosomal RNA; rRNA; bacterial infection; human; food grain supplement; livestock; poultry; therapeutic; ds.

XX OS Vibrio cholerae.

XX PN WO20142457-A2.

XX PD 14-JUN-2001.

XX PF 29-NOV-2000; 2000WO-US042391.

XX PR 29-NOV-1999; 99US-016150P.

XX PA (AVIB-) AVI BIOPHARMA INC.

XX PI Iversen PL;

XX DR WPI; 2001-457295/49.

XX PT Antibacterial compound, useful for treating bacterial infections and as

PT livestock and poultry food supplement, comprises antisense

PT oligonucleotides complementary to bacterial 16S and 23S rRNA.

XX PS Disclosure; Page; 62PP; English.

XX CC AAS11021-AAS11034 represent the coding sequences of bacterial 16S

CC ribosomal RNA (rRNA) genes. The sequences were used to design anti-  
CC bacterial compounds comprising substantially unchanged antisense

CC oligomer containing 8-40 nucleotide subunits, including a targeting

CC nucleic acid sequence at least 10 nucleotides in length which is

CC complementary to a bacterial 16S or 23S rRNA nucleic acid sequence. The

CC antisense oligomers are used for treating a bacterial infection in a

human or a mammalian animal produced by *Escherichia coli*, *Salmonella typhimurium*, *Pseudomonas aeruginosa*, *Vibrio cholera*, *Neisseria gonorrhoea*, *Helicobacter pylori*, *Bartonella henselae*, *Hemophilus influenzae*, *Shigella dysenteriae*, *Staphylococcus aureus*, *Mycobacterium tuberculosis*, *Streptococcus pneumoniae*, *Treponema pallidum* and *Chlamydia trachomatis*. The antibacterial compound may be used as a food grain supplement in livestock and poultry feed composition. Note: The present sequence is not shown in the specification but has been accessed from Genbank using the appropriate accession number given in the specification. (Updated on 11-SEP-2003 to standardise OS field)

XX SQ Sequence 1500 BP; 376 A; 326 C; 482 G; 312 T; 0 U; 4 Other;

Query Match 64.0%; Score 19.2; DB 5; Length 1500;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02; Mismatches 2; Indels 5; Gaps 0;

Qy 1 CTGGACCCCTCTGACTCGAGTCGCT 28  
Db 212 CTGGGCCATCCCGACGCCGARAGTCCG 185

RESULT 3

ADP97107/c  
ID ADP97107 standard; DNA; 121124 BP.

XX AC ADP97107;

XX DT 07-OCT-2004 {first entry}

DE Mouse cancer associated sequence MD08-002, SEQ ID 83.

XX KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.

OS Mus musculus.

XX RN WO2004-060304-A2.

XX PD 22-JUL-2004.

XX PF 22-DEC-2003; 2003WO-US041389.

XX PR 27-DEC-2002; 2002US-00330773.

XX PA (SAGR-) SAGRES DISCOVERY INC.

XX PI Morris DW, Malandro MS;

XX DR WPI; 2004-543781/52.

XX PT New isolated cancer associated nucleic acids comprising at least 10 contiguous nucleotides, useful for diagnosing, preventing and/or treating cancers such as leukemia and lymphoma.

XX PS Claim 1; SEQ ID NO 83; 199pp; English.

The present invention relates to cancer associated sequences (ADQ97025-ADQ98004). The sequences are useful for the diagnosis, prevention and/or treatment of cancer, such as leukemia and lymphoma. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 121124 BP; 32972 A; 25314 C; 26641 G; 35451 T; 0 U; 746 Other;

Query Match 63.3%; Score 19; DB 12; Length 121124;  
Best Local Similarity 81.5%; Pred. No. 1.9e+02; Mismatches 22; Indels 5; Gaps 0;

Qy 4 GACCCCTCTGACTCGAGTCGCT 30  
Db 88764 GCCGCCATCCCGACGCCGARAGTCCG 88738

RESULT 4

XX	Organization
SQ	Sequence 2481 BP; 599 A; 653 C; 555 G; 674 T; 0 U; 0 Other;
XX	Query Match Similarity 62.7%; Score 18.8; DB 10; Length 2481;
AC	Best Local 76.7%; Pred. No. 1.7e-02; Mismatches 7; Indels 0; Gaps 0;
XX	Matches 23; Conservative 0; MisMatches 7; Indels 0; Gaps 0;
DT	26-JUN-2003 (first entry)
DE	Toxicity modelling related rat gene SEQ ID No 1468.
XX	KW toxic effect; gene expression profile; renal toxicity; toxicity marker; database; drug screening; toxicity assay; rat; ds.
KW	OS Rattus norvegicus.
XX	PN WO200295000-A2.
XX	PD 28-NOV-2002.
XX	PF 22-MAY-2002; 2002WO-US016173.
XX	PR 22-MAY-2001; 2001US-0292335P.
PR	13-JUN-2001; 2001US-0297523P.
PR	19-JUN-2001; 2001US-0298025P.
PR	10-JUL-2001; 2001US-030307P.
PR	10-JUL-2001; 2001US-030308P.
PR	10-JUL-2001; 2001US-0303010P.
PR	28-AUG-2001; 2001US-031047P.
PR	27-SEP-2001; 2001US-0324928P.
PR	22-OCT-2001; 2001US-0330462P.
PR	01-NOV-2001; 2001US-0330567P.
PR	21-NOV-2001; 2001US-0331805P.
PR	01-DEC-2001; 2001US-0336144P.
PR	19-DEC-2001; 2001US-03340873P.
PR	21-FEB-2002; 2002US-0357042P.
PR	21-FEB-2002; 2002US-0357043P.
PR	21-FEB-2002; 2002US-0357044P.
PR	15-MAR-2002; 2002US-0361349P.
PR	01-APR-2002; 2002US-0370144P.
PR	08-APR-2002; 2002US-0370206P.
PR	08-APR-2002; 2002US-0370247P.
PR	17-APR-2002; 2002US-0372794P.
PR	21-APR-2002; 2002US-0371679P.
PA	(GENE-) GENE LOGIC INC.
XX	PI Mendick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX	DR WPI; 2003-148464/14.
XX	PT Predicting at least one toxic effect of a compound, useful for toxicity modeling, comprises preparing a gene expression profile of a tissue or cell sample exposed to the compound, and comparing the gene expression profile to a database.
PS	Example 4; Page: 446pp; English.
XX	The invention relates to a novel method of predicting at least one toxic effect of a compound. The method comprises a gene expression profile of a tissue or cell sample exposed to the compound, and comparing the gene expression profile to a database comprising at least part of the data or information given in the specification. The methods are useful for predicting at least one toxic effect of a compound, predicting the progression of a toxic effect of a compound, predicting the renal toxicity of a compound, or identifying toxicity markers in tissues or cells exposed to known renal toxin. The genes are useful as toxicity markers in drug screening and toxicity assays, in monitoring disease or physiological states or disease progression. This polynucleotide represents a rat DNA sequence relating to the toxic effect database described in the specification. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the World Intellectual Property Organization.

RESULT 5

XX	Organization
SQ	Sequence 23107 BP; 5308 A; 6264 C; 6229 G; 5306 T; 0 U; 0 Other;
XX	Query Match Similarity 62.7%; Score 18.8; DB 10; Length 23107;
AC	Best Local 76.7%; Pred. No. 1.7e-02; Mismatches 7; Indels 0; Gaps 0;
XX	Matches 23; Conservative 0; MisMatches 7; Indels 0; Gaps 0;
Qy	1 CTGGACCCCTCTGACTCGAGAGTCGCT 30
Db	1077 CTGAGCCCTCTGCTAGGGAGTTCACT 1048

RESULT 6

XX	Organization
ID	ADA02762 standard; DNA; 23107 BP.
XX	AC ADA02762;
XX	PD 06-NOV-2003 (first entry)
XX	DT 06-NOV-2003 (first entry)
DB	Human RUNX3 carcinoma associated gene, SEQ ID No:1280.
XX	KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast; prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening; gene; ds.
OS	Homo sapiens.
XX	PN WO2003057146-A2.
XX	PR 17-JUL-2003.
XX	PD 17-JUL-2003.
XX	PR 26-DEC-2002; 2002WO-US041414.
XX	PR 26-DEC-2001; 2001US-00035832.
XX	(SAGR-) SAGRES DISCOVERY.
XX	PI Morris DW;
XX	DR WPI; 2003-587068/55.
XX	PT New recombinant nucleic acid encoding carcinoma associated protein, useful for preparing compositions for treating carcinomas.
XX	PS Claim 1; SEQ ID NO 1280; 245pp; English.
XX	The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADK0182-ADK0304), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a bichip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.
SQ	Sequence 23107 BP; 5308 A; 6264 C; 6229 G; 5306 T; 0 U; 0 Other;

Query Match 62.7%; Score 18.8; DB 9; Length 23107;  
 Best Local Similarity 76.7%; Pred. No. 2.1e+02; Mismatches 7; Indels 0; Gaps 0;  
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTGACTCGAGAGTCGGT 30  
 Db 14873 CTGGAGGCCCTGGAATTCGGT 14902

RESULT 6

ID ADB2500 standard; DNA; 23107 BP.

XX ADB2500;  
 AC XX  
 DT 04-DEC-2003 (first entry)

DE Human Runx3 gene.

XX human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;  
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.  
 OS Homo sapiens.

XX WO2003008583-A2.

XX 30-JAN-2003.

XX 26-DEC-2001; 2001WO-US051291.

XX 02-MAR-2001; 2001US-0079586.

PR 23-OCT-2001; 2001US-00004113.

PR 08-NOV-2001; 2001US-00052482.

PR 30-NOV-2001; 2001US-00997722.

PR 20-DEC-2001; 2001US-00034650.

PA (SAGR-) SAGRES DISCOVERY.

XX Morris DW, Engelhard EK;

XX WPI; 2003-513603/48.

PT New recombinant nucleic acid comprising a nucleotide sequence of any of  
 PT the carcinoma-associated (CA) genes, useful for screening for drug  
 PT candidates for diagnosing or treating carcinomas.

XX Claim 1; SEQ ID NO 28; 983pp; English.

CC The invention relates to a recombinant nucleic acid comprising a  
 CC nucleotide sequence selected from any of the fully defined carcinoma-  
 CC associated (CA) genes from the 50 tables given in the specification. The  
 CC proteins are secreted, transmembrane or intracellular proteins. The  
 CC recombinant nucleic acids are useful for screening for drug candidates  
 CC for diagnosing or treating carcinomas. Sequences given in ADC8514-  
 CC represent CA genes of the invention.

SQ Sequence 23107 BP; 5308 A; 6264 C; 5306 T; 0 U; 0 Other;

Query Match 62.7%; Score 18.8; DB 10; Length 23107;  
 Best Local Similarity 76.7%; Pred. No. 2.1e+02; Mismatches 7; Indels 0; Gaps 0;  
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTGACTCGAGAGTCGGT 30  
 Db 14873 CTGGAGGCCCTGGAATTCGGT 14902

RESULT 8

ID ADM74357 standard; DNA; 23107 BP.

XX ADM74357;  
 AC XX  
 DT 01-JUL-2004 (first entry)

DE Human carcinoma associated (CA) nucleic acid #13.

XX Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;  
 KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;  
 KW cytostatic.

XX Homo sapiens.

XX US2004072154-A1.

XX 15-APR-2004.

XX 30-NOV-2001; 2001US-00997722.

PR 22-DEC-2000; 2000US-00741377.

PR 02-MAR-2001; 2001US-0079586.

PA (MORR/) MORRIS D W.

RESULT 7

ID ADC85242 standard; DNA; 23107 BP.

XX ADC85242;  
 AC XX  
 DT 01-JAN-2004 (first entry)

PA (ENGE/) ENGELHARD B K.  
 XX  
 PI Morris DW, Engelhard EK;  
 XX  
 DR WPI; 2004-328562/30.  
 XX  
 PT New carcinoma associated gene or protein, useful for preparing a composition for diagnosing or treating carcinoma e.g., leukemia or lymphoma.  
 XX  
 PS Claim 1; SEQ ID NO 28; 29pp; English.  
 XX  
 CC The invention relates to new recombinant nucleic acids. The invention also relates to a host cell comprising a recombinant nucleic acid or expression vector, an expression vector comprising a recombinant nucleic acid, a recombinant protein, a method of screening for drug candidates, a method of screening for a biocative agent capable of binding to a carcinoma associated protein (CAP) encoded by a nucleotide sequence, a method of screening for a biocative agent capable of modulating the activity of a CAP, a method of evaluating the effect of a candidate carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting the activity of a CAP, a method of treating carcinomas, a method of neutralising the effect of a CAP and a method of diagnosing carcinoma or propensity to carcinoma. A method of evaluating the effect of a candidate carcinoma drug comprises administering the drug to a patient, removing a cell sample from the patient and determining alterations in the expression or activation of a gene comprising the nucleotide sequence. A method of diagnosing carcinoma comprises determining the expression of one or more genes comprising the nucleic acid sequence in a first tissue type of a first individual and comparing the expression of the gene from a second normal tissue type from the first individual or a second unaffected individual, where a difference in the expression indicates the first individual has carcinoma. A method of inhibiting the activity of a CAP comprises binding an inhibitor to the CAP. Treating carcinomas comprises administering to a patient an inhibitor of CAP. Neutralising the effect of a CAP comprises contacting an agent specific for the CAP. The polypeptide specifically binds to the protein encoded by the nucleic acid. It comprises an antibody that specifically binds to the protein encoded by the nucleic acid. The nucleic acids are useful for preparing a composition for diagnosing or treating carcinoma e.g., leukaemia or lymphoma. This sequence represents a human carcinoma associated (CA) nucleic acid of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.  
 CC  
 XX Sequence 23107 BP; 5308 A; 6264 C; 6229 G; 5306 T; 0 U; 0 Other;  
 CC  
 CC Best Local Similarity 62.7%; Score 18.8; DB 12; Length 23107;  
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 QC 1 CTGGACCCCTCTGACTCGAGAGTTGCCT 30  
 CC  
 Db 14902 CTGGGAGCCCCCTGGACTCCAGATTCCGCT 14902  
 CC  
 SQ Sequence 3210 BP; 1050 A; 626 C; 595 G; 939 T; 0 U; 0 Other;  
 CC  
 CC Best Local Similarity 62.0%; Score 18.6; DB 2; Length 3210;  
 CC Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 CC  
 QC 1 CTGGACCCCTCTGACTCGAGAGTT 25  
 CC  
 Db 1232 CTGGACCCCTCTGAAATGGATAGTT 1208  
 CC  
 SQ Sequence 3210 BP; 1050 A; 626 C; 595 G; 939 T; 0 U; 0 Other;  
 CC  
 CC Best Local Similarity 62.0%; Score 18.6; DB 2; Length 3210;  
 CC Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 CC  
 QC 1 CTGGACCCCTCTGACTCGAGAGTT 25  
 CC  
 Db 1232 CTGGACCCCTCTGAAATGGATAGTT 1208  
 CC  
 RESULT 9  
 ARZ2/624/C  
 ID ARZ2/624 standard; DNA; 3210 BP.  
 XX  
 AC ARZ2/624;  
 XX  
 DT 20-DEC-1999 (first entry)  
 XX  
 DE Plasmid SPP-1.  
 XX  
 KW Extracellular compartment modification; floral cell; self-compatibility;  
 KW pollen-pistil interaction; self-incompatibility; insect growth control;  
 KW Plasmid SPP-1; GPI363 gene; cysteine protease inhibitor gene; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO949063-A1.  
 XX  
 PD 30-SEP-1999.  
 XX  
 PP 19-MAR-1999; 99WO-CA000237.  
 XX  
 PR 20-MAR-1998; 98US-0078728P.  
 XX  
 PA (MIAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.  
 XX  
 PI Robert LS, Gleddie S;  
 XX  
 DR WPI; 1999-591104/50.  
 XX  
 CC Protein expression in floral cells for peptide display, mediating plant sterility, and modifying pollen-pistil interactions.  
 XX  
 PS Example 6; Page 103-104; 113pp; English.  
 XX  
 CC This sequence represents the plasmid SP1F-1, containing a fusion of the Brassica napus GPI363 gene to the Onchocerca protease inhibitor gene. The invention relates to a method for modifying the extracellular compartment of a floral cell of a plant, that comprises expressing a construct comprising a gene of interest encoding a protein, fusion protein or peptide, or a fragment of them, which is capable of modifying the composition of the extracellular compartment of the floral cell and altering either the function, use or development of the floral cell or modifying the interaction of the floral cell with other cells, within an anther or pistil cell. The method is used to modify pollen-pistil interaction or function, which mediates, produces or prevents self-compatibility, self-incompatibility, out- or in-crossing or combinations of these. The method is also used for localizing proteins on the surface of pollen for the purpose of peptide display. The protein localized on the surface of the pollen may be an antibody or antigen or is a protein that is effective in controlling insect growth, behaviour, feeding, development or reproduction.  
 CC  
 XX Sequence 3210 BP; 1050 A; 626 C; 595 G; 939 T; 0 U; 0 Other;  
 CC  
 CC Best Local Similarity 62.0%; Score 18.6; DB 2; Length 3210;  
 CC Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 CC  
 QC 1 CTGGACCCCTCTGACTCGAGAGTT 25  
 CC  
 Db 1232 CTGGACCCCTCTGAAATGGATAGTT 1208  
 CC  
 RESULT 10  
 ARZ2/623/C  
 ID ARZ2/623 standard; DNA; 3459 BP.  
 XX  
 AC ARZ2/623;  
 XX  
 DT 20-DEC-1999 (first entry)  
 XX  
 DE Plasmid SPP-1.  
 XX  
 KW Extracellular compartment modification; floral cell; self-compatibility;  
 KW pollen-pistil interaction; self-incompatibility; insect growth control;  
 KW Plasmid SPP-1; GPI363 gene; cysteine protease gene; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO949063-A1.  
 XX  
 PD 30-SEP-1999.  
 XX  
 PP 19-MAR-1999; 99WO-CA000237.  
 XX  
 PR 20-MAR-1998; 98US-0078728P.  
 XX  
 PA (MIAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.

XX  
PI Robert LS, Gleddie S;  
XX DR  
DR WI; 1999-591104/50.

PT Protein expression in floral cells for peptide display, mediating plant  
sterility, and modifying pollen-pistil interactions.

PS Example 5; Page 100-102; 113pp; English.

XX  
CC This sequence represents the plasmid SPF-1, containing a fusion of the  
Brassica napus GPR33 gene to the sitophilus cysteine protease gene. The  
invention relates to a method for modifying the extracellular compartment  
of a floral cell of a plant, that comprises expressing a construct  
comprising a gene of interest encoding a protein, fusion protein or  
peptide, or a fragment of them, which is capable of modifying the  
composition of the extracellular compartment of the floral cell and  
altering either the function, use or development of the floral cell or  
modifying the interaction of the floral cell with other cells, within an  
anther or pistil cell. The method is used to modify pollen-pistil  
interaction or function, which mediates, produces or prevents self-  
compatibility, self-incompatibility, out- or in-crossing or combinations  
of these. The method is also used for localizing proteins on the surface  
of pollen for the purpose of peptide display. The protein localized on  
the surface of the pollen may be an antibody or antigen or is a protein  
that is effective in controlling insect growth, behaviour, feeding,  
development or reproduction

SQ Sequence 3459 BP; 1082 A; 721 C; 678 G; 978 T; 0 U; 0 Other;

Query Match 62.0%; Score 18.6; DB 2; Length 3459;  
Best Local Similarity 84.0%; Pred. No. 2.2e-02; Matches 21;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTCGACTCGAGGTT 25  
Db 1232 CTGACCCCTTCGATGGATGT 1208

RESULT 11

ID AAH51601 standard; DNA; 319608 BP.

AC AAH51601;

XX DT 29-AUG-2001 (first entry)

DB Human chromosome 13q31-q33 genomic nucleotide sequence.  
KW sbg1; g34665; sbg2; g35017; g35018; chromosome 13q31-q33; haplotype;  
KW biallelic marker; polymorphism; schizophrenia; bipolar disorder; ds.  
OS Homo sapiens.

XX PN WO2000058510-A2.  
XX PD 05-OCT-2000.  
XX DT 30-MAR-2000; 2000WO-IB00435.

PR 30-MAR-1999; 99US-0126903P.  
PR 30-APR-1999; 99US-0131971P.  
PR 14-JUL-1999; 99US-0132055P.  
PR 27-JUL-1999; 99US-0143920P.  
PR 29-JUL-1999; 99US-0145915P.  
PR 29-JUN-1999; 99US-0146525P.  
PR 28-OCT-1999; 99US-0162288P.  
XX PA (GEST ) GENSET.

XX PI Cohen D, Blumenfeld M, Chumakov I, Bouquelert L, Bihain B;

PI Bessieux L;

XX WPI; 2000-619082/59.

XX DR  
XX Polynucleotides comprising sequences from sbg1 and g35018 biallelic  
PT markers are used for genotyping and detecting schizophrenia or bipolar  
PT disorder and predisposition to these disorders.

PS Claim 1; Page 409-493; 737pp; English.

XX  
CC AAH51601 represents a human genomic nucleotide sequence comprising sbg1,  
g3665, sbg2, g3017 and g35018 nucleic acid sequences located on the  
human chromosome 13q31-q33 locus. The nucleotide sequences contain  
biallelic markers and polymorphisms. Sequences AAH51602 - AAH51626 and  
AAH62907 - AAH62915 represent cDNA human sbg1 cDNA sequences and protein  
products. AAH51627 - AAH5631 and AAH62916 - AAH62918 represent g35018  
cDNA sequences and protein products. Primers AAH51632 - AAH51699 are used  
to isolate sbg1 cDNAs, while sbg1 exons from different primates are  
represented by sequences AAH51642 - AAH51699. Nucleotide sequences of  
amplicons which comprise biallelic markers located on the chromosome  
13q31-q33 locus are represented in AAH5170 - AAH5187. Biallelic markers  
are represented in the sequences by degenerate/undefined base codes. PCR  
primers AAH51818 and AAH5199 are used in the isolation of sequences of  
the invention. The biallelic marker containing nucleotide sequences are  
used to determine the identity of the nucleotide at a biallelic marker in  
a sample DNA sequence. The nucleotide sequences may be labelled and used  
for genotyping by determining the identity of a nucleotide at a Region D-  
related biallelic marker in a biological sample from single or multiple  
subjects. By determining the frequency of a biallelic marker in a  
population an association between a genotype and a trait, a haplotype and  
a trait and a phenotype and a trait can be detected. The sequences can be  
used to determine a predisposition to or early onset of schizophrenia or  
bipolar disorder or a beneficial response to or side effects related to  
treatment against schizophrenia or bipolar disorder

XX  
CC SQ Sequence 319608 BP; 101600 A; 56677 C; 58335 G; 102722 T; 0 U; 274 Other;  
Query Match 61.3%; Score 18.4; DB 3; Length 319608;  
Best Local Similarity 78.6%; Pred. No. 3.8e+02; Matches 22;  
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTGGACCCCTCTCGACTCGAGAGTCCG 28  
Db 8690 CTGAACCATCTCGATTGAGATTCAG 8717

RESULT 12

ID AAH50301 standard; DNA; 319608 BP.

AC AAH50301;

XX DT 26-SEP-2001 (first entry)

DB Human schizophrenia associated gene g35030 and biallelic markers A1-A71.

XX KW Human; g35030; biallelic marker; A1-A71; chromosome 13q31-q33;

XX OS Homo sapiens.

XX PR 05-OCT-2000.

XX DT 30-MAR-2000; 2000WO-IB00435.

PR 30-APR-1999; 99US-0131971P.  
PR 14-JUL-1999; 99US-0132055P.  
PR 27-JUL-1999; 99US-0143920P.  
PR 29-JUL-1999; 99US-0145915P.  
PR 29-JUN-1999; 99US-0146525P.  
PR 28-OCT-1999; 99US-0162288P.

XX PA (GEST ) GENSET.

XX PI Cohen D, Blumenfeld M, Chumakov I, Bouquelert L, Bihain B;

PI Bessieux L;

XX  
CC FT Key  
FT primer\_bind location/Qualifiers  
FT /\*tag= a  
FT /note= "Binds primer 99-27943.rp"  
FT primer\_bind 8297. .8315 /\*tag= b  
FT /note= "Binds primer 99-27943-150.mis"  
FT misc\_binding 8304. .8328 /\*tag= c  
FT /bind\_moiety= "Probe\_99-27943-150"  
FT misc\_feature 8316 /\*tag= d  
FT /note= "Biallelic marker A1"

```

FT primer_bind complement(8317. .8335)
FT /*tag= e
FT /notes= "Binds primer 99-27943-150.mis complement"
FT complement(8446. .8465)
FT /*tag= f
FT /notes= "Binds primer 99-27943.pu complement"
FT 21365. .21385
FT /*tag= g
FT /notes= "Binds primer 99-27935.rP"
FT 21653. .21671
FT /*tag= h
FT /notes= "Binds primer 99-27935-193.mis"
FT 21660. .21684
FT /*tag= i
FT /bound_moiety= "Probe_99-27935-193"
FT misc_feature 21672
FT /*tag= j
FT /notes= "Biallelic marker A2"
FT primer_bind complement(21673. .21691)
FT /*tag= k
FT /notes= "Binds primer 99-27935-193.mis complement"
FT complement(21845. .21864)
FT /*tag= l
FT /notes= "Binds primer 99-27935.pu complement"
FT 65463. .65471
FT /*tag= m
FT /notes= "Binds primer 8-128.pu"
FT primer_bind 65465. .65484
FT /*tag= n
FT /notes= "Binds primer 8-128-33.mis"
FT misc_binding 65473. .65497
FT /*tag= o
FT /bound_moiety= "Probe_8-128-33"
FT misc_feature 65485
FT /*tag= p
FT /notes= "Biallelic marker A3"
FT primer_bind complement(65486. .65504)
FT /*tag= q
FT /notes= "Binds primer 8-128-33.mis complement"
FT primer_bind complement(65656. .65874)
FT /*tag= r
FT /notes= "Binds primer 8-128.rP complement"
FT primer_bind 95034. .95053
FT /*tag= s
FT /notes= "Binds primer 99-31960.pu"
FT primer_bind 95377. .95395
FT /*tag= t
FT /notes= "Binds primer 99-31960-363.mis"
FT misc_binding 95384. .95408
FT /*tag= u
FT /bound_moiety= "Probe_99-31960-363"
FT misc_feature 95395
FT /*tag= v
FT /notes= "Biallelic marker A4"
FT primer_bind complement(95397. .95415)
FT /*tag= w
FT /notes= "Binds primer 99-31960-363.mis complement"
FT primer_bind complement(95543. .95563)
FT /*tag= x
FT /notes= "Binds primer 99-31960.rP complement"
FT primer_bind 107022. .107040
FT /*tag= y
FT /notes= "Binds primer 99-24656.pu"
FT primer_bind 107252. .107280
FT /*tag= z
FT /notes= "Binds primer 99-24656-260.mis"
FT misc_binding 107259. .107293
FT /*tag= aa
FT /bound_moiety= "Probe_99-24656-260"
FT misc_feature 107281
FT /*tag= ab
FT /notes= "Biallelic marker A5"
FT complement(107282. .107300)

FT primer_bind /*tag= ac
FT /notes= "Binds primer 99-24656-260.mis complement"
FT complement(107495. .107513)
FT /*tag= ad
FT /notes= "Binds primer 99-24656.rP complement"
FT primer_bind 160279. .160298
FT /*tag= ae
FT /notes= "Binds primer 99-24639.rP"
FT primer_bind 160631. .160639
FT /*tag= af
FT /notes= "Binds primer 99-24639-163.mis"
FT misc_binding 160628. .160652
FT /*tag= ag
FT /bound_moiety= "Probe_99-24639-163"
FT misc_feature 160640. .160678
FT /*tag= ah
FT /notes= "Biallelic marker A6"
FT primer_bind 160641. .160659
FT /*tag= ai
FT /notes= "Binds primer 99-24639-163.mis complement"
FT primer_bind 160770. .160787
FT /*tag= aj
FT /notes= "Binds primer 99-24634.pu"
FT primer_bind complement(160785. .160802)
FT /*tag= ak
FT /notes= "Binds primer 99-24634.pu complement"
FT primer_bind 160857. .160875
FT /*tag= al
FT /notes= "Binds primer 99-24634-108.mis"
FT misc_binding 160864. .160888
FT /*tag= am
FT /bound_moiety= "Probe_99-24634-108"
FT misc_feature 160876
FT /*tag= an
FT /notes= "Biallelic marker A7"
FT primer_bind 160888
FT /*tag= ao
FT /notes= "Binds primer 99-24634-108.mis complement"
FT primer_bind complement(161240. .161254)
FT /*tag= ap
FT /notes= "Binds primer 99-24634.rP complement"
FT primer_bind 168913. .168930
FT /*tag= ar
FT /notes= "Binds primer 99-24634.rP complement"
FT misc_binding 168932. .168936
FT /*tag= as
FT /bound_moiety= "Probe_99-7652-162"
FT misc_feature 168974
FT /*tag= at
FT /notes= "Biallelic marker A8"
FT primer_bind complement(168975. .168993)
FT /*tag= au
FT /notes= "Binds primer 99-7652-162.mis complement"
FT primer_bind complement(169331. .169351)
FT /*tag= av
FT /notes= "Binds primer 99-7652.rP complement"
FT primer_bind 170666. .170686
FT /*tag= aw
FT /notes= "Binds primer 99-16100.pu"
FT primer_bind 170791. .170809
FT /*tag= ax
FT /notes= "Binds primer 99-16100-147.mis"
FT misc_binding 170798. .170822
FT /*tag= ay
FT /bound_moiety= "Probe_99-16100-147"
FT misc_feature 170810
FT /*tag= az
FT /notes= "Biallelic marker A9"
FT primer_bind complement(170811. .170829)
FT /*tag= ba

```



```

PT /*tag= ae
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= af
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT replace(20491, T)
PT /*tag= ag
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= ah
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT replace(22248, G)
PT /*tag= ai
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= aj
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT replace(23789, A)
PT /*tag= ak
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= al
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT replace(23930, A)
PT /*tag= am
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= an
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT replace(26339, A)
PT /*tag= ao
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= ap
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= ar
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= at
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= au
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= av
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= ax
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= ay
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= az
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT /*tag= bb
PT /standard name= "Single nucleotide polymorphism"
PT variation
PT replace(33671, A)
PT /*tag= bc

```

---

```

FT variation
FT /standard name= "Single nucleotide polymorphism"
FT replace(35045, C)
FT /*tag= bd
FT /standard name= "Single nucleotide polymorphism"
FT variation
FT replace(35944, G)
FT /*tag= be
FT /standard name= "Single nucleotide polymorphism"
FT replace(37157, G)
FT /*tag= bf
FT /standard name= "Single nucleotide polymorphism"
FT variation
FT replace(37603, T)
FT /*tag= bg
FT /standard name= "Single nucleotide polymorphism"
FT replace(39242, G)
FT /*tag= bh
FT /standard name= "Single nucleotide polymorphism"
FT replace(39404, A)
FT /*tag= bi
FT /standard name= "Single nucleotide polymorphism"
FT replace(40395, C)
FT /*tag= bj

```

Query Match Similarity 60.7%; Score 18.7; DB 8; Length 340449;  
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 25999 TGGACCCCTCCACAGKT 26021

RESULT 14  
 ADD74382 DT 04-DEC-2003 (first entry)  
 ID ADD74382 standard; DNA; 42325 BP.  
 XX AC ADB74382;  
 XX DT 04-DEC-2003 (first entry)  
 DB Mycobacterium leprae DNA #16.  
 XX KW Non-naturally occurring peptide; anion pump protein; tuberculosis;  
 XX hypersensitivity reaction; tuberculostatic; gene; ds.  
 OS Mycobacterium leprae.  
 XX PR 19-AUG-1993; 93US-00109181.  
 PR 22-OCT-1993; 93US-00142558.  
 XX PN US653266-B1.  
 XX PD 24-JUN-2003.  
 XX PP 16-SEP-1994; 94US-00311731.  
 XX PR 19-AUG-1993; 93US-00109181.  
 PR 22-OCT-1993; 93US-00142558.  
 XX PA (GENO-) GENOME THERAPEUTICS CORP.  
 PI Smith DR, Mao J;  
 XX DR WPI; 2003-656441/62.  
 XX New Mycobacterium tuberculosis anion pump peptide useful for as  
 PT tuberculosis vaccine and diagnosis of tuberculosis infection.  
 PT Disclosure; SEQ ID NO 131; 26pp; English.

CC The invention relates to a non-naturally occurring peptide of  
 CC Mycobacterium tuberculosis comprising an amino acid sequence  
 CC corresponding to an anion pump protein. The invention also relates to a  
 CC non-naturally occurring nucleic acid corresponding to a DNA sequence of  
 CC Mycobacterium tuberculosis or Mycobacterium leprae. The new peptide is  
 CC useful as a vaccine against Mycobacterium tuberculosis or Mycobacterium  
 CC leprae or for screening for new tuberculosis drugs. Purified proteins

CC derived from the sequences of the invention may elicit a specific immune response. The peptide may also be used to detect hypersensitivity reactions of individuals exposed to Mycobacterium tuberculosis or Mycobacterium leprae. The proteins and peptides may be affixed to solid supports to detect antibodies typical of hypersensitivity reactions, from a patient's sera. This sequence represents Mycobacterium leprae DNA of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 42325 BP; 9673 A; 13128 C; 11330 G; 8194 T; 0 U; 0 Other;

Query Match 60.0%; Score 18; DB 10; Length 42325;  
Best Local Similarity 80.8%; Pred. No. 5e+02; Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 ACCCCUCTCGACTCGAGAGTCGGT 30  
Db 688 ACCCCACTCGAACAGTGACGCT 713

RESULT 15

ADM27081\_10/c Continuation (11 of 17) of ADM27081 from base 1000001 (Hyperthermophile Methanopyrus kandleri)

WP Sequence	Fragment Name	Begin	End	LOCUS	Accession
WP	ADM27081_00	1	110000	ADM27081	
WP	ADM27081_01	100001	210000		
WP	ADM27081_02	200001	310000		
WP	ADM27081_03	300001	410000		
WP	ADM27081_04	400001	510000		
WP	ADM27081_05	500001	610000		
WP	ADM27081_06	600001	710000		
WP	ADM27081_07	700001	810000		
WP	ADM27081_08	800001	910000		
WP	ADM27081_09	900001	1010000		
WP	ADM27081_10	1000001	1110000		
WP	ADM27081_11	1100001	1210000		
WP	ADM27081_12	1200001	1310000		
WP	ADM27081_13	1300001	1410000		
WP	ADM27081_14	1400001	1510000		
WP	ADM27081_15	1500001	1610000		
WP	ADM27081_16	1600001	1694968		

Query Match 60.0%; Score 18; DB 11; Length 110000;  
Best Local Similarity 80.8%; Pred. No. 5.4e+02; Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CTGGACCCCTCTCCTCGACTCGAGAGTC 26  
Db 69684 CTTGGCCGCTCTGACTCGAGAGTC 69659

Search completed: February 28, 2005, 23:51:44  
Job time : 268.818 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: February 28, 2005, 23:32:36 ; Search time 875.065 Seconds  
(without alignments)  
1661.198 Million cell updates/sec

Title: US-10-086-062-4

Perfect score: 30

Sequence: 1 ctggacccttcgtactcgagatccgct 30

Scoring table: IDENTITY\_NUC

Gapext 1.0 , Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422/607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

GenBank,\*

1: gb\_bp:\*

2: gb\_hhg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ovr:\*

6: gb\_prt:\*

7: gb\_ph:\*

8: gb\_pr:\*

9: gb\_pr:\*

10: gb\_r0:\*

11: gb\_sb:\*

12: gb\_sv:\*

13: gb\_un:\*

14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description	ORIGIN	ALIGMENTS
1	30	100.0	30	AX358109	RESULT 1		
2	20.4	68.0	595	9 HUMYQ76F01	AX358109	LOCUS	AX358109
3	20.4	68.0	98348	9 AL136103	Sequence 4 from Patent WO01943394.	DEFINITION	30 bp
4	20.4	68.0	17865	9 AC012048	AC012048	ACCESSION	PAT 13-FEB-2002
5	20.2	67.3	23211	3 AC099767	AC099767	VERSION	
6	20	66.7	1503	1 AB038027	AB038027	KEYWORDS	GI:18674856
7	20	66.7	156165	2 AC021453	AC021453	SOURCE	
8	20	66.7	157980	9 AC00852	AC00852	ORGANISM	
9	19.6	65.5	257959	2 AC123011	AC123011	REFERENCE	
10	19.4	64.7	11406	4 AF07201	AF07201	AUTHORS	Jilka, J.M., Hood, E.B. and Howard, J.A.
11	19.4	64.7	70887	3 AC087075	AC087075	TITLE	Novel plant promoter sequences and methods of use for same
12	19.4	64.7	110000	8 AB016816	AB016816	JOURNAL	Patent: WO 01943394-A 4 13-DEC-2001; Prodigene, Inc. (US)
13	19.4	64.7	159669	9 AC0104791	AC0104791	FEATURES	Location/Qualifiers
14	19.4	64.7	189669	2 AF006440	AF006440	1. .30	
15	19.4	64.7	196772	2 AC011862	AC011862	/mol type="unassigned DNA"	
16	19.2	64.0	1500	1 AF118021	AF118021	/db_xref="taxon:358"	
17	19.2	64.0	1500	6 AR052360	AR052360		
18	19.2	64.0	1500	6 AX201179	AX201179		
19	63.3	93695	8 ATPT2009	ATPT2009	ATPT2009	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 595) Woesener,J., Tan,J., Marra,M., Kucaba,T., Yandell,M., Martin,J., Marth,G., Bowles,L., Wyllie,T., Bowers,Y., Steptoe,M., Theising,B., Geisel,S., Allen,M., Underwood,K., Chappell,J., Person,B., Gibbons,M., Harvey,N., Pape,D., Chamberlain,A., Morales,R., Schurk,R., Ritter,E., Kohn,S., Swaller,T., Behymer,K., Hillier,L., Wilson,R. and Waterston,R.
TITLE	Full Clone Sequencing of the Longest Available Member from Each Unigene Cluster
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 595) Waterson,R.
AUTHORS	Direct Submission
TITLE	Submitted (24-AUG-1998) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT	Submitted By: Genome Sequencing Center Department of Genetics Washington University St. Louis MO 63108, USA <a href="http://genome.wustl.edu/gsc">http://genome.wustl.edu/gsc</a> mailto:eat@watson.wustl.edu
FEATURES	The location of this clone is unknown. Location/Qualifiers
source	1. .595 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:201721" /clone_lib="Soares_fetal_liver_spleen_INFSL" 9. .338 /rpt_family="L2" repeat_region
ORIGIN	repeat_region
Query Match	Best local similarity 80.0%; Score 20.4; DB 9; Length 595; Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy	1 CTGACCCCTCTGACTCGAGTTCGGT 30
Db	566 CTGAGCCCTCTCGCTGAGACTTCCT 537
RESULT	3
AL136103	AL136103
DEFINITION	Human DNA sequence from clone RP1-250B11 on chromosome 10 Contains STS and GSS, complete sequence.
ACCESSION	AL136103
VERSION	AL136103.24
KEYWORDS	HTG
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 98348) Philimore,B.
AUTHORS	Direct Submission
TITLE	Submitted (04-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
COMMENT	CB10 1SA, UK. E-mail enquiries: <a href="mailto:hunqury@sanger.ac.uk">hunqury@sanger.ac.uk</a> Clone requests: <a href="mailto:clone.request@sanger.ac.uk">clone.request@sanger.ac.uk</a> On Jul 20, 2000 this sequence version replaced gi:950151. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Mp:, WORMPEP; Information on the WORMPEP database can be found at <a href="http://www.sanger.ac.uk/Projects/C_elegans/wormpep">http://www.sanger.ac.uk/Projects/C_elegans/wormpep</a> This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <a href="http://www.sanger.ac.uk/HGP/Chro10">http://www.sanger.ac.uk/HGP/Chro10</a> Pieter de Jong. For further details see <a href="http://www.chori.org/bacpac/home.htm">http://www.chori.org/bacpac/home.htm</a>
FEATURES	Location/Qualifiers
source	1. .98348 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="10" /clone="RP1-250B11" /clone_lib="RPCI-1" /clone_id="RPCI-1" /misc_feature
repeat_region	1. .34 /note="Single clone region. Reads derived from clone PCR. Assembly consistent with restriction digest." 2458-.765 /note="Alusx repeat: matches 1..308 of consensus" repeat_region
repeat_region	/note="MER39b repeat: matches 355..546 of consensus" 3249..3606 /note="MER39 repeat: matches 13..380 of consensus" 6106..6337 /note="L2 repeat: matches 2435..2705 of consensus" repeat_region
repeat_region	/note="Alusg repeat: matches 1..298 of consensus" complement(8333..8779) /note="match: GSS: Em:AQ403443" /note="L2 repeat: matches 2435..2705 of consensus" repeat_region
misc_feature	/note="match: GSS: Em:AQ264393" 8302..9234 /note="match: GSS: Em: AQL05662" misc_feature
repeat_region	/note="match: GSS: Em: AQL05662" 9268..9356 /note="match: STS: Em: G28019" repeat_region
repeat_region	11809..12092 /note="Alusx repeat: matches 1..284 of consensus" repeat_region
repeat_region	12205..12222 /note="L1ME repeat: matches 5285..5811 of consensus" repeat_region
repeat_region	14113..14171 /note="L2 repeat: matches 2692..2750 of consensus" 14548..14846 /note="AluY repeat: matches 1..299 of consensus" repeat_region
repeat_region	15883..16169 /note="AlusX repeat: matches 1..287 of consensus" 17146..18557 /note="MER52A repeat: matches 1..1755 of consensus" 18559..18713 /note="MER repeat: matches 6..160 of consensus" repeat_region
repeat_region	18955..19381 /note="MER63 repeat: matches 590..1061 of consensus"

---

repeat_region	19383..19519	repeat_region	49784..49912
/note="FLAM C repeat: matches 3..137 of consensus"		/note="FLAM A repeat: matches 8..132 of consensus"	
repeat_region	19713..19878	repeat_region	49914..50164
/note="MEB63 repeat: matches 463..577 of consensus"		/note="MIR repeat: matches 2..261 of consensus"	
repeat_region	19826..20117	repeat_region	52850..52918
/note="MEB63 repeat: matches 2..307 of consensus"		/note="L2 repeat: matches 2635..2706 of consensus"	
repeat_region	20325..20391	misc_feature	53092..53613
/note="MIR repeat: matches 88..156 of consensus"		misc_feature	/note="match: GSS: Em:AQ800506"
repeat_region	20440..20701	misc_feature	/note="match: GSS: Em:AQ172450"
/note="MIR repeat: matches 2..248 of consensus"		misc_feature	53151..53604
repeat_region	22321..22631	misc_feature	/note="match: GSS: Em:AQ170181"
/note="match: STS: Em:HSL10C4"		misc_feature	53165..53586
repeat_region	25728..26188	misc_feature	/note="match: GSS: Em:AQ211002"
/note="match: GSS: Em:AQ211002"		misc_feature	53179..53355
repeat_region	25738..25978	misc_feature	/note="match: GSS: Em:AQ898390"
/note="match: GSS: Em:AQ823705"		repeat_region	53376..53834
repeat_region	26548..27067	misc_feature	/note="match: GSS: Em:AQ898390"
/note="match: GSS: Em:AQ98605"		repeat_region	53976..53958
repeat_region	27824..28952	misc_feature	/note="MIR repeat: matches 41..250 of consensus"
/note="L2 repeat: matches 2167..2416 of consensus"		repeat_region	54076..54203
repeat_region	28725..29750	misc_feature	/note="match: GSS: Em:AQ055423"
/note="match: GSS: Em:AQ055423"		repeat_region	54204..54216
repeat_region	29759..29793	misc_feature	/note="MIR repeat: matches 250..260 of consensus"
/note="match: GSS: Em:B87793"		repeat_region	54316..54767
repeat_region	29852..30037	misc_feature	/note="match: GSS: Em:AQ155973"
/note="match: GSS: Em:AQ155973"		repeat_region	54379..54565
repeat_region	30073..30670	misc_feature	/note="match: GSS: Em:AQ729256"
/note="match: GSS: Em:AQ195052"		repeat_region	54532..54565
repeat_region	30670..31212	misc_feature	/note="MIR repeat: matches 65..308 of consensus"
/note="match: GSS: Em:AQ099305"		repeat_region	54440..54730
repeat_region	31212..32344	misc_feature	/note="match: GSS: Em:AQ084956"
/note="match: GSS: Em:AQ859233"		repeat_region	54731..55355
repeat_region	32344..33368	misc_feature	/note="match: GSS: Em:AQ556262"
/note="match: GSS: Em:AQ24293"		repeat_region	55610..55935
repeat_region	34554..34583	misc_feature	/note="match: GSS: Em:AQ594663"
/note="match: GSS: Em:AQ223744"		repeat_region	56871..57293
repeat_region	35059..36135	repeat_region	/note="L1MD repeat: matches 883..1319 of consensus"
/note="L2 repeat: matches 2417..2500 of consensus"		repeat_region	57237..57602
repeat_region	36191..36191	/note="AluJo repeat: matches 1..305 of consensus"	
/note="match: GSS: Em:B55923"		repeat_region	
repeat_region	36191..36192	repeat_region	
/note="match: GSS: Em:AQ03876"		repeat_region	
repeat_region	36692..36800	repeat_region	
/note="MIR repeat: matches 13..123 of consensus"		repeat_region	
repeat_region	40461..40765	repeat_region	
/note="ALUSX repeat: matches 1..293 of consensus"		repeat_region	
repeat_region	41312..41502	repeat_region	
/note="MIR repeat: matches 8..192 of consensus"		repeat_region	
repeat_region	41785..41840	repeat_region	
/note="28 copies 2 mer tg 85% conserved"		repeat_region	
repeat_region	43022..43334	repeat_region	
/note="ALUSX repeat: matches 3..312 of consensus"		repeat_region	
repeat_region	43447..45066	repeat_region	
/note="match: GSS: Em:AQ540223"		repeat_region	
repeat_region	44671..45063	repeat_region	
/note="match: GSS: Em:AQ337658"		repeat_region	
repeat_region	45077..45157	repeat_region	
/note="MEB57 internal repeat: matches 7151..7230 of		repeat_region	
consensus"		repeat_region	
repeat_region	45158..45304	repeat_region	
/note="LIP repeat: matches 5050..5197 of consensus"		repeat_region	
repeat_region	45284..45621	repeat_region	
/note="LIP repeat: matches 5476..5812 of consensus"		repeat_region	
repeat_region	45911..46452	repeat_region	
/note="match: GSS: Em:AQ032143"		repeat_region	
repeat_region	48564..48709	repeat_region	
/note="L2 repeat: matches 2596..2749 of consensus"		repeat_region	
repeat_region	48720..49082	repeat_region	
/note="THE1B repeat: matches 1..364 of consensus"		repeat_region	
repeat_region	49175..49369	repeat_region	
/note="MLTID repeat: matches 1..502 of consensus"		repeat_region	

---

		TITLE	Direct Submission
		JOURNAL	Submitted (16-AUG-2001) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA
REFERENCE	4	AUTHORS	Smith, D.R.
		TITLE	Direct Submission
		JOURNAL	Submitted (27-MAR-2002) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA
REFERENCE	5	AUTHORS	Smith, D.R.
		TITLE	Direct Submission
		JOURNAL	Submitted (28-MAR-2002) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA
COMMENT	On Mar 27, 2002 this sequence version replaced gi:15193325.	FEATURES	(bases 1 to 178965)
FEATURES	source	SOURCE	Smith, D.R.
			Direct Submission
			Submitted (20-APR-2002) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA
COMMENT	On Mar 27, 2002 this sequence version replaced gi:15193325.	FEATURES	(bases 1 to 178965)
		SOURCE	Location/Qualifiers
			1..178965
			/organism="Homo sapiens"
			/mol_type="genomic DNA"
			/db_xref="taxon:9606"
			/chromosome="10"
			/clone="RP11-33N22"
			/clone_1lb="RCI-11"
		ORIGIN	
RESULT 5		Query Match	Score 68.0%; Score 20.4; DB 9; Length 178965;
AC093767/c		LOCUS	Best Local Similarity 80.0%; Pred. No. 1.9e+02; Mismatches 6; Indels 0; Gaps 0;
		DEFINITION	Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY	1	ACCESSION	GI:17017639
		KEYWORDS	
Db	32913	VERSION	CTGAGCCCTCTCGACTGGAGTCCT 30
		ORGANISM	Caenorhabditis briggsae
		REFERENCE	
		AUTHORS	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae; Rhabditidae; Pelederinae; Caenorhabditis.
		TITLE	Caenorhabditis briggsae
		JOURNAL	Washington University Genome Sequencing Center.
		REFERENCE	The C. briggsae Genome Sequencing Project
		AUTHORS	Unpublished
		TITLE	2 (base 1 to 23211)
		JOURNAL	Waterson, R.
		REFERENCE	Direct Submission
		AUTHORS	Unpublished
		TITLE	3 (bases 1 to 23211)
		JOURNAL	Waterson, R.
		REFERENCE	Direct Submission
		AUTHORS	Submitted (20-NOV-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
COMMENT	Submitted by: Genome Sequencing Center	FEATURES	
		SOURCE	Department of Genetics, Washington University St. Louis , MO 63110, USA email: rwm@nematoe.wustl.edu
NOTICE:	This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.	ORIGIN	
RESULT 6		Query Match	Score 67.3%; Score 20.2; DB 3; Length 23211;
AB038027/c		LOCUS	Best Local Similarity 88.0%; Pred. No. 2.7e+02; Mismatches 3; Indels 0; Gaps 0;
		DEFINITION	Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	4	ACCESSION	GI:7007473
		KEYWORDS	
Db	14462	VERSION	GACCCCTCTGATCGAGAGTTCG 28
		ORGANISM	Vibrio sp. SG128
		REFERENCE	
		AUTHORS	Bacterial; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibionaceae; Vibrio.
		TITLE	16S rRNA gene of marine bacterium
		JOURNAL	Published Only in DataBase (2000)
		REFERENCE	2 (bases 1 to 1503)
		AUTHORS	Urakawa, H.
		TITLE	Direct Submission
		JOURNAL	Submitted (05-FEB-2000) Hideyoshi Urakawa, Northwestern University, Department of Civil Engineering; Technological Institute 2145 Sheridan Road, Evanston, Illinois 60208-3109, USA (E-mail:h-urakawa@nwu.edu, Tel:+1-847-467-5710, Fax:+1-847-491-4011) Location/Qualifiers
		FEATURES	
		SOURCE	1..1503
			/organism="Vibrio sp. SG128"
			/mol_type="genomic DNA"
			/strain="SG128"
			/db_xref="taxon:115126"
			<1..1503>
			/product="16S ribosomal RNA"
RESULT 7		Query Match	Score 66.7%; Score 20; DB 1; Length 1503;
AC021453/c		LOCUS	Best Local Similarity 82.1%; Pred. No. 4.1e+02; Mismatches 5; Indels 0; Gaps 0;
		DEFINITION	Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY	1	ACCESSION	GI:7362318
		VERSION	CTGGACCCCTCTCGACTGGAGTTCCG 28
Db	231	VERSION	CTGGACCCCTCTCGACTGGAGTTCCG 204

**KEYWORDS**  
 HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
**SOURCE**  
 Homo sapiens (human)  
**ORGANISM**  
 Homo sapiens  
 Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE**  
 1 (bases 1 to 156165)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
**AUTHORS**  
 Homo sapiens, clone RPI1-125C16  
**JOURNAL**  
 Unpublished  
**REFERENCE**  
 2 (bases 1 to 156165)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,  
 Boguslavsky, L., Bouknight, B., Brown, A., Burkett, G., Castle, A.,  
 Choepl, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
 DeArleano, K., Dekker, K., Domino, M., Doyle, M., Fenestor, J.,  
 Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,  
 Gardina, S., Grant, G., Higgs, B., Howland, J. C., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Howland, J. C., Jones, R., Jones, C., Kaufman, R., Klein, J.,  
 Landers, T., Lebockzy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,  
 Macdonald, P., Marquis, N., McEwan, P., McCurk, A., McKernan, K.,  
 MacPheevers, R., Meldrim, J., Meneau, L., Morrow, J., Nayler, J.,  
 Norman, C.H., O'Connor, T., O'Donnell, P.P., Oliver, T.M., Peterson, K.,  
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
 Roy, A., Santos, R., Severy, B., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Subramanian, A., Talama, J., Testave, S., Theodore, J.,  
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,  
 Zimmer, A. and Zody, M.  
**TITLE**  
 Direct Submission  
**JOURNAL**  
 Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**COMMENT**  
 On Apr 1, 2000 this sequence version replaced g1.6721267.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
**FEATURES**  
**SOURCES**  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone\_id="RPI1-125C16"  
 /clone\_lib="RPCI-11 Human Male BAC"  
**1..1836**  
*/note="assembly\_fragment"*  
*1937..5796*  
*/note="assembly\_fragment"*  
*5891..9797*  
*/note="assembly\_fragment"*  
*9898..13919*  
*/note="assembly\_fragment"*  
*14000..17400*  
*/note="assembly\_fragment"*  
*17501..21253*  
*/note="assembly\_fragment"*  
*clone\_end:SP5*  
*vector\_end:T7*  
*vector\_end:right"*  
*vector\_end:T7*  
*vector\_end:right"*  
*89975..116362*  
*/note="assembly\_fragment"*  
*11663..156165*  
*/note="assembly\_fragment"*  
**LOCUS**  
 AC10052/c  
 AC100852  
 LOCUS AC100852  
 DEFINITION Homo sapiens chromosome 17, clone RPI1-125C16,  
 ACCESSION AC100852  
 VERSION AC100852.2  
 KEYWORDS HTG, HTGS\_PHASE1, HTGS\_DRAFT  
**ORIGIN**  
 Query Match 66.7%; Score 20; DB 2; length 157980 bp  
 Best Local Similarity 82.1%; Pred. No. 2.9e-02; Mismatches 5; Indels 1  
 Matches 23; Conservative 0; MisMatche 5; InDel 1  
**RESULT**  
 Oy 2 TGGACCCCTCTGAGCTGAGGTTTCGC 29  
 Db 11285 TGGACCTCTCTTCAAGGTTCTGC 11258

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammali; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 ('bases 1 to 15780')
JOURNAL	Unpublished
TITLE	Homo sapiens chromosome 17, clone RP11-125C16
2 ('bases 1 to 15780')	
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguskiy, L., Boukhalter, B., Brown, A., Camarillo, J., Castellano, K., Collins, S., Collymore, A., Cook, A., Choopel, Y., Colangelo, M., Coppiiano, A., Chang, J., Chazaro, B., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, W., Gage, D., Galagan, J., Gardina, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Haigis, B., Headford, A., Horton, L., Huime, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Kakatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lebozky, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Melorim, J., Menous, L., Mihova, T., Milenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Poliara, V., Raymond, C., Reetta, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spender, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tessyave, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Vilim, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
JOURNAL	Direct Submission
REFERENCE	Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 330 Charles Street, Cambridge, MA 02141, USA
AUTHORS	3 ('bases 1 to 157980')
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Basien, V., Blouin, T., Boguskiy, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choopel, Y., Collymore, A., Cook, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardina, S., Gord, S., Grand-Pierre, N., Haigis, B., Horton, L., Huime, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Mihova, T., Milenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Reetta, R., Roy, A., Seaver, P., Rogov, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Vilei, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
JOURNAL	Direct Submission
COMMENT	Submitted (29-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 330 Charles Street, Cambridge, MA 02141, USA On Aug 29, 2002 this sequence version replaced g1:17048222. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <a href="http://ftp.genome.washington.edu/RM/RepeatMasker.html">http://ftp.genome.washington.edu/RM/RepeatMasker.html</a>
FEATURES	<p>Center project name: L21587</p> <p>Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: <a href="http://www-seq.wi.mit.edu">http://www-seq.wi.mit.edu</a> Contact: sequence submissions@genome.wi.mit.edu</p> <p>Center project name: L21587</p> <p>Center clone name: 125_C_16</p> <p>Location/Qualifiers</p> <p>1. .157980</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="genomic DNA"</p>

	AUTHORS	Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angliano, D., Aryal, E., Balcic, V., Aoyagi, M., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bjorn, O., Blahay, C., Burch, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burrell, J., Calderon, E., Cardenas, V., Carter, K., Carvazo, J., Ceasar, H., Centner, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Andrade, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandes, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisa, B., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebrgeorgis, B., Geer, K., Gill, R., Grady, M., Guera, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Harwey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodges, A., Hogues, M., Hollins, B., Howell, S., Huilyk, S., Hume, J., Idlebrian, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Lorenshewa, L., Louisged, H., Lozano, R.J., Lu, X., Ma, J., Maneshwari, M., Manindartne, M., Mahmood, M., Malloy, K., Mangum, A., Manjum, B., Mapua, P., Martin, K., Martin, R., Martinez, B., Mawhiney, S., McLeod, M.P., McNeil, T.Z., Meenan, E., Milosavljevic, A., Miner, G., Minia, E., Montemayor, J.J., Moore, S., Morgan, M., Morris, K., Morris, M., Mundaca, S., Muniz, S., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakelume, O., Okwuonu, G., Olakunmabogun, A., Pal, S., Parks, K., Pasternak, S., Perez, P., Perez, L., Pfankoch, C., Plopper, R., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Schefer, S., Scott, G., Shatsman, S., Shen, H., Shettty, J., Shvartsburg, A., Sisson, I., Sitter, C.D., Smajsić, S., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleczik, R., Woeden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yatub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinsrock, G., and Gibbs, R.A.	JOURNAL
RESULT	TITLE	JOURNAL	
Query Match	66.7%; Score 20; DB 9; Length 157980;	JOURNAL	
Best Local Similarity	82.1%; Pred. No. 2.9e+02;	UNPUBLISHED	
Matches	23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;		
Qy	2 TGACGACCTCTGACTCGAGAGTTCCGC 29		
Db	130676 TGGACCTCTCTTATCAAGAGTTCTGC 130649		
RESULT	TITLE	JOURNAL	
AC123011	257595 bp DNA linear HTG 12-OCT-2002	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (26-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		
SOURCE	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 257595)		
RESULT	TITLE	JOURNAL	
AC123011	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	JOURNAL	
DEFINITION	Rattus norvegicus clone CH230-92124, WORKING DRAFT SEQUENCE, 4	UNPUBLISHED	
ACCESSION	unordered pieces.		
VERSION	AC123011		
KEYWORDS	AC123011..3 GI:23665280		

genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine

Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

----- Project Information  
Center project name: GYWD  
Center clone name: CH230-92124

----- Summary Statistics

Assembly program: prap; version 0.990329

Consensus quality: 236342 bases at least Q40

Consensus quality: 239870 bases at least Q30

Consensus quality: 242005 bases at least Q20

Estimated insert size: 243933; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a "working draft" sequence. It currently

\* consists of 4 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 250677: contig of 250677 bp in length

\* 250678 250777: gap of unknown length

\* 250778 252307: contig of 15330 bp in length

\* 252308 252407: gap of unknown length

\* 252408 253577: contig of 1170 bp in length

\* 253578 253677: gap of unknown length

\* 253678 257595: contig of 3918 bp in length.

## FEATURES

### source

1. 257595 /organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-92124"

/contig="246393; 247143"

/note="clone boundary"

clone\_end:Sp6

end\_sequence:BH295371"

/note="clone boundary"

clone\_end:Sp6

2 (bases 1 to 11406)

3 (bases 1 to 236342)

4 (bases 1 to 239870)

5 (bases 1 to 242005)

6 (bases 1 to 243933)

7 (bases 1 to 250677)

8 (bases 1 to 250678)

9 (bases 1 to 250777)

10 (bases 1 to 252307)

11 (bases 1 to 252407)

12 (bases 1 to 253577)

13 (bases 1 to 253677)

14 (bases 1 to 257595)

REFERENCE	Mammalia; Eutheria; Perissodactyla; Equidae; Equus. 1 (bases 1 to 11406)
AUTHORS	Masel,A.M., Brandon,R.B. and Bell,T.K.
TITLE	Nucleotide Sequence of the Equine Beta-Lactoglobulin Gene
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 11406)
AUTHORS	Masel,A.M., Brandon,R.B. and Bell,T.K.
TITLE	Direct Submission
JOURNAL	Submitted (17-Nov-1998) Australian Equine Blood Typing Research Laboratory, University of Queensland, St. Lucia, Brisbane, Queensland 4072, Australia

Location/Qualifiers  
1. .11406 /organism="Equus caballus"  
/mol\_type="genomic DNA"  
/db\_xref="ttaxon:9796"  
<4472..>6679  
>6679  
/gene="BLG"  
/join(<4472..4567,4956..5095,5767..5840,7367..7480,  
8302..8406,8663..8679)  
/gene="BLG"  
/product="beta-lactoglobulin II"  
/join(4472..4567,4956..5095,5767..5840,7367..7480,  
8302..8406,8663..8679)  
/codon\_start=1  
/product="beta-lactoglobulin II"  
/protein\_id="AAC93385.1"  
/db\_xref="IGI:008112"  
/translation="MVKCLALISLMCGNQADIPQMDDIDLQEAVGRWASVAMV  
SDISLDSSESPFLRYVBLRPTPEGNITIRREGANHACVVRNTIVQKTERPAVTV  
NYQERKSVLTDYAHMFVCFQGPPLPSAHEGMWCQLARTQKVDEMEFRSLQ  
PLPGRVQIVQDPSQQERGF"

## ORIGIN

### Query Match

### Best Local Similarity

### Matches

### Definition

### Accession

### Version

### Keywords

### Source

### Organism

### Reference

### Authors

### Title

### Journal

## RESULT

### 11

### AC087075

### DEFINITION

### Caenorhabditis briggsae cosmid CB023K10, complete sequence.

### ACCESSION

### AC087075

### VERSION

### 5.1

### GI:

### 11545986

### KEYWORDS

### HTG.

### SOURCE

### Caenorhabditis briggsae

### ORGANISM

### Eukaryota; Metazoa; Hematoda; Chromadorea; Rhabditida;

### Rhabditidae; Rhabditidae; Peloderrinae; Caenorhabditis.

### REFERENCE

### 1 (bases 1 to 70387)

### AUTHORS

### Washington University Genome Sequencing Center.

### TITLE

### The C. briggsae Genome Sequencing Project

### JOURNAL

### Unpublished

### REFERENCE

### 2 (bases 1 to 70387)

### AUTHORS

### Peterson,R.

## COMMENT

### Submitted by:

### Genome Sequencing Center

### Department of Genetics, Washington University,

### St. Louis, MO 63110, USA

### e-mail: jspieh@watson.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections.

once, or longer because we provide a small overlap between neighboring submissions.

#### JOURNAL/Qualifiers

FEATURES  
Source

I. .70387

/organism="Caenorhabditis briggsae"

/mol\_type="genomic DNA"

/db\_xref="Gujarat\_G16"

/db\_xref="taxon:6238"

/clones="CBP023K10"

23138. .23209

/product="tRNA-Gly"

/notes="codon recognized: GGC"

complement(36599..36640)

/product="tRNA-Glu"

/notes="codon recognized: GAG"

37437. .37508

/product="tRNA-Gln"

/notes="codon recognized: CAA"

ORIGIN

Query Match

Best Local Similarity

Matched

Score 64.7%;

DB 3;

Length 70387;

Pred. No. 5.7e+02;

Mismatches 0;

Indels 6;

Gaps 0;

Fragment Name

Begin

End

AE016816\_0

1

110000

AE016816\_1

100001

210000

AE016816\_2

200001

310000

AE016816\_3

300001

410000

AE016816\_4

400001

510000

AE016816\_5

500001

610000

AE016816\_6

600001

710000

AE016816\_7

700001

810000

AE016816\_8

800001

907057

Continuation (7 of 9) of AE016816 from base 600001 (AE016816 *Bremothecium gossypii* ATCC

Query Match

Best Local Similarity

Matched

Score 64.7%;

DB 8;

Length 11000;

Pred. No. 5.5e+02;

Mismatches 0;

Indels 6;

Gaps 0;

QY

1

CCTGGACCCCTCTGACTCGAGGTCCGC

29

Db

20976

CTGGACTCTGTGAATGAGCTTCCAC

20948

RESULT 13

AC104791/c

AC104791

159969 bp

DNA

linear

PRI 21-FEB-2002

DEFINITION

Homo sapiens BAC clone RP11-18IK12 from 4, complete sequence.

ACCESSION

AC104791 AC032008

VERSION

AC104791\_3

G1:1842313

KEYWORDS

HTG,

HTG, Homo sapiens

Homo sapiens (human),

Homo sapiens

Homo sapiens (human),

REFERENCE

1. (bases 1 to 15969)

Sulston,J.E. and Watson,R.

Toward a complete human genome sequence

Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE

99053792

PUBLISHED

1994074

REFERENCE

2. (bases 1 to 159969)

Isak,A., Meyer,R. and Creason,K.

TITLE The sequence of *Homo sapiens* BAC clone RP11-18IK12  
JOURNAL Unpublished (2001)  
REFERENCE 3 (bases 1 to 159969)  
AUTHORS Waterston, R.H.

REFERENCE JOURNAL Submitted (21-DEC-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 159969)

AUTHORS Waterston, R.H.

REFERENCE JOURNAL Direct Submission Submitted (03-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 159969)

AUTHORS Waterston, R.

REFERENCE JOURNAL Direct Submission Submitted (21-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

COMMENT On Feb 3, 2002 this sequence version replaced gi:18042374.

----- Genome Center Center: Washington University Genome Sequencing Center

Web site: http://genome.wustl.edu

Contact: baptiens@wustl.edu

----- Summary Statistics Center project name: H\_NH018IK12

Drafting Center: WIBR

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:**  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

**SOURCE INFORMATION:**  
The RP11-1 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Cataneo,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong VECTOR: pBACe3.6

**NEIGHBORING SEQUENCE INFORMATION:**  
The clone sequenced to the left is RP11-229M9; the clone sequenced to the right is RP11-203B7. Actual start of this clone is at base position 1 of RP11-18IK12; actual end is at base position 159969 of RP11-18IK12.

A transposon has been identified in the vector of this clone.

The sequence of AC032008 has been incorporated into AC104791.  
FEATURES source  
REFERENCE 1..159969  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

```

/chromosome="4"
/map="4"
/clone="RP11-181K12"
/clone_lib="RPCI-11"
191. .463
repeat_region
/rept_family="MIR"
1855. .2217
/rept_family="MAlR"
2333. .2362
repeat_region
/rept_family="AT_rich"
2471. .2690
repeat_region
/rept_family="MIR"
3143. .3169
/rept_family="(A)n"
3296. .3316
repeat_region
/rept_family="AT_rich"
3304. .3362
/rept_family="Alu"
3589. .3609
/rept_family="AT_rich"
3984. .4033
repeat_region
/rept_family="BRLV"
4019. .4111
/rept_family="(TA)n"
4087. .4495
/rept_family="BRLV"
5159. .5483
/rept_family="Alu"
5374. .5402
/rept_family="(GGATG)n"
5897. .5927
/rept_family="(TTTTCG)n"
5906. .6217
/rept_family="Alu"
6577. .6760
/note="similar to Homo sapiens EST AW843469
(NID:gr937452)" .7131
/note="similar to Homo sapiens EST AW843469
(NID:gr937452)" .7314. .7393
/rept_family="MIR"
/rpt_family="(T)n"
7700. .7766
repeat_region
/rept_family="MIR"
9154. .9438
/rept_family="Alu"
9437. .9466
/rept_family="(TAA)n"
10074. .10204
repeat_region
/rept_family="MIR"
10351. .10812
repeat_region
/rept_family="MAlR"
11128. .11343
repeat_region
/rept_family="Alu"
11414. .11626
repeat_region
/rpt_family="MERL_type"
13354. .13449
repeat_region
/rpt_family="MERL_type"
13450. .13721
repeat_region
/rpt_family="Alu"
13722. .13784
repeat_region
/rpt_family="MERL_type"
13813. .13852
repeat_region
/rpt_family="CT-rich"
15896. .15941
repeat_region
/rpt_family="(TGC)n"
16404. .16916
repeat_region
/rpt_family="L2"
16918. .16939
repeat_region
/rpt_family="(T)n"
18098. .18220

repeat_region
/rept_family="L1"
18250. .18339
repeat_region
/rept_family="Alu"
18545. .18594
repeat_region
/rept_family="(A)n"
18829. .19005
/note="match to EST AA743144 (NID:92782650) ny20h09.s1"
repeat_region
/rept_family="Alu"
19565. .19869
repeat_region
/rept_family="Alu"
20046. .20785
repeat_region
/rept_family="ERV1"
21452. .21689
repeat_region
/rept_family="L1"
21764. .22061
repeat_region
/rept_family="Alu"
21780. .22172
misc_feature
/note="match to EST AA554485 (NID:92325024) ni36b03.s1"
repeat_region
/rept_family="Alu"
22691. .22990
repeat_region
/rept_family="Alu"
23268. .23312
repeat_region
/rept_family="(TTA)n"
23286. .23596
repeat_region
/rept_family="Alu"
23637. .24312
repeat_region
/rept_family="Alu"
24463. .24756
repeat_region
/rept_family="L2"
24795. .25023
repeat_region
/rept_family="MER2_type"
25050. .25411
repeat_region
/rept_family="BRLV"
25377. .25406
repeat_region
/rept_family="(CAAA)n"
25412. .25733
repeat_region
/rept_family="Alu"
25532. .25561
/rept_family="(A)n"

Query Match
Best Local Similarity 64.7%; Score 194; DB 9; Length 159969;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Ov
Db 120906 TGGACTRACTCTGCTCCAGACTCTCT 120878

RESULT 14
AP006440/c
LOCUS AP006440
DEFINITION Homo sapiens chromosome 11 clone RP11-368120 map 11q, WORKING DRAFT
SEQUENCE 4 unordered pieces.
ACCESSION AP006440
VERSION AP006440.1 GI:3096286
KEYWORDS HTG; HGCS; PHASEL; HTGS_DRAFT.
SOURCE HTG; HGCS; PHASEL; HTGS_DRAFT.
ORGANISM Homo sapiens (human)
REFERENCE Hattori, M., Ishii, K., Toyota, A., Taylor, T.D., Hong-Seqg, P., Makiyama, A., Yada, T., Tokoi, Y., Watanabe, H. and Sakai, Y.
AUTHORS Hattori, M., Ishii, K., Toyota, A., Taylor, T.D., Hong-Seqg, P., Makiyama, A., Yada, T., Tokoi, Y., Watanabe, H. and Sakai, Y.
TITLE Homo sapiens genomic DNA of 11q
JOURNAL Published Only in Database (2003)
REFERENCE 2 (bases 1 to 18969)
AUTHORS Hattori, M., Ishii, K., Toyota, A., Taylor, T.D., Hong-Seqg, P., Makiyama, A., Yada, T., Tokoi, Y., Watanabe, H. and Sakai, Y.
TITLE Direct Submission
JOURNAL Submitted (19-MAY-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Saito-cho, Tsukuba, Ibaraki-ku, Yokohama, Kanagawa 230-0045, Japan
E-mail:hattori@gsc.riken.go.jp; URL: http://hgpc.gsc.riken.go.jp/; Tel: 81-45-503-9111, Fax: 81-45-503-9170

```

COMMENT	LOCUS	AC011862
	DEFINITION	Homo sapiens clone RP11-15K8, LOW-PASS SEQUENCE SAMPLING.
	ACCESSION	AC011862
	VERSION	AC011862.4 GI:7144913
	KEYWORDS	HTG; HTGS; PHASE0.
-----	SOURCE	Homo sapiens
Center project name: HumanRP11	ORGANISM	Homo sapiens (human)
Center clone name: RP11-368120	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
-----	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
Sequencing vector: PCR products; 100% of reads	JOURNAL	Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Assembly program: Phrap; version 0.99029	REFERENCE	Unpublished
Chemistry: Dye-terminator ET-amerham; 100% of reads	AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Consensus quality: 188057 bases at least Q40	TITLE	1 (bases 1 to 186472)
Consensus quality: 188482 bases at least Q30	JOURNAL	Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Insert size: 188969; sum-of-contigs	REFERENCE	Baldwin, J., Barna, N., Beckerly, R., Boguskiy, L., Boukhalter, B.,
Quality coverage: 12.20x in Q20 bases; sum-of-contigs	AUTHORS	Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
-----	ORGANISM	Cooke, P., Dearellano, K., Dewart, J., Domino, M., Donelan, L., Doyle, M.,
NOTE: This is a 'working draft' sequence. It currently consists of	REFERENCE	Ferreira, P., FitzHugh, M., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyne, S., Grant, G., Hedges, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Medrano, J., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Strange-Thomann, N., Stoyanovic, N., Subramanian, A., Talamas, J., Teffaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
4 contigs. The true order of the pieces is not known and their	REFERENCE	Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charle Street, Cambridge, MA 02141, USA
order in this sequence record is arbitrary. Gaps between the	AUTHORS	Teufay, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
contigs are represented as runs N, but the exact sizes of the gaps	TITLE	All repeats were identified using RepeatMasker:
are unknown. This record will be updated with the finished sequence	JOURNAL	Smit, A. F. A. & Green, P. (1996-1997) <a href="http://ftp.genome.washington.edu/RM/RepeatMasker.html">http://ftp.genome.washington.edu/RM/RepeatMasker.html</a>
as soon as it is available and the accession number will be	COMMENT	-----
preserved.	COMMENT	Center: Whitehead Institute/ MIT Center for Genome Research
-----	REFERENCE	Center code: WIBR
1. 70266 contig of 70266 bp in length	AUTHORS	Web site: <a href="http://www-sq.wi.mit.edu">http://www-sq.wi.mit.edu</a>
* 70367 70366: gap of 100 bp	TITLE	Contact: <a href="mailto:sequence_submissions@genome.wi.mit.edu">sequence_submissions@genome.wi.mit.edu</a>
* 70367 130873: contig of 60507 bp in length	JOURNAL	-----
* 130874 130973: gap of 100 bp	COMMENT	Project Information
* 130974 188629: contig of 57556 bp in length	REFERENCE	Center project name: 15_K_8
* 188630 188729: gap of 100 bp	AUTHORS	-----
* 188730 189269: contig of 540 bp in length.	TITLE	-----
FEATURES	COMMENT	-----
Source	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT	-----
-----	REFERENCE	-----
-----	AUTHORS	-----
-----	TITLE	-----
-----	JOURNAL	-----
-----	COMMENT</	

8855 9749: contig of 895 bp in length \* 45234 45333: gap of 100 bp  
 9750 9849: gap of 100 bp \* 45334 46202: contig of 869 bp in length  
 9850 10738: contig of 889 bp in length \* 46203 46302: gap of 100 bp  
 10739 10838: gap of 100 bp \* 46303 47193: contig of 891 bp in length  
 10839 11725: contig of 887 bp in length \* 47194 47293: gap of 100 bp  
 11726 11825: gap of 100 bp \* 47294 48182: contig of 889 bp in length  
 11826 12707: contig of 882 bp in length \* 48183 48282: gap of 100 bp  
 12708 12807: gap of 100 bp \* 48283 49146: contig of 864 bp in length  
 12808 13697: contig of 890 bp in length \* 49147 49246: gap of 100 bp  
 13698 13797: gap of 100 bp \* 49247 50116: contig of 870 bp in length  
 13798 14668: contig of 871 bp in length \* 50117 50216: gap of 100 bp  
 14669 14768: gap of 100 bp \* 50217 51112: contig of 896 bp in length  
 14769 15648: contig of 880 bp in length \* 51113 51212: gap of 100 bp  
 15649 15748: gap of 100 bp \* 51213 52102: contig of 890 bp in length  
 15749 16623: gap of 100 bp \* 52103 52202: gap of 100 bp  
 16623 17587: contig of 865 bp in length \* 52203 53064: contig of 862 bp in length  
 17588 17687: gap of 100 bp \* 53065 53164: gap of 100 bp  
 17688 18558: contig of 871 bp in length \* 53165 54024: contig of 860 bp  
 18559 18558: gap of 100 bp \* 54025 54124: gap of 100 bp  
 18659 19534: contig of 874 bp in length \* 54126 54937: contig of 833 bp in length  
 19535 19634: gap of 100 bp \* 54937 55057: gap of 100 bp  
 19635 20510: contig of 876 bp in length \* 55058 55939: contig of 881 bp in length  
 20511 20610: gap of 100 bp \* 55939 56028: gap of 100 bp  
 20611 21317: contig of 871 bp in length \* 56029 56932: contig of 860 bp  
 21318 21617: gap of 100 bp \* 56932 57032: gap of 100 bp  
 21618 22499: contig of 882 bp in length \* 57032 57916: contig of 884 bp in length  
 22499 22500: gap of 100 bp \* 57916 58016: gap of 100 bp  
 22500 22599: contig of 876 bp in length \* 58016 58017: contig of 859 bp in length  
 22599 22599: gap of 100 bp \* 58017 58875: contig of 881 bp in length  
 22600 23480: contig of 881 bp in length \* 58876 58976: contig of 864 bp in length  
 23481 23580: gap of 100 bp \* 58976 59845: gap of 100 bp  
 23581 24444: contig of 864 bp in length \* 59845 59944: gap of 100 bp  
 24445 24444: gap of 100 bp \* 59945 59945: contig of 851 bp in length  
 24545 25479: contig of 935 bp in length \* 59945 60795: gap of 100 bp  
 25479 25579: gap of 100 bp \* 60795 60895: contig of 879 bp in length  
 25579 25580: contig of 913 bp in length \* 60895 61774: contig of 879 bp in length  
 25580 25592: gap of 100 bp \* 61774 61874: gap of 100 bp  
 25592 27454: contig of 862 bp in length \* 61874 62739: contig of 865 bp in length  
 27454 27454: gap of 100 bp \* 62739 62839: gap of 100 bp  
 27555 28431: contig of 877 bp in length \* 62839 63807: contig of 968 bp in length  
 28431 28531: gap of 100 bp \* 63807 63907: gap of 100 bp  
 28532 28532: contig of 964 bp in length \* 63907 64783: contig of 876 bp in length  
 28532 29495: contig of 913 bp in length \* 64783 64883: gap of 100 bp  
 29495 29595: gap of 100 bp \* 64883 65767: contig of 884 bp in length  
 29595 30490: contig of 894 bp in length \* 65767 65867: gap of 100 bp  
 30490 30590: gap of 100 bp \* 65867 66765: contig of 898 bp in length  
 30590 31435: contig of 846 bp in length \* 66765 66868: gap of 100 bp  
 31436 31535: gap of 100 bp \* 66868 66866: contig of 865 bp in length  
 31536 32425: contig of 890 bp in length \* 66866 67730: gap of 100 bp  
 32426 32525: gap of 100 bp \* 67730 67731: contig of 891 bp in length  
 32525 33401: contig of 876 bp in length \* 67731 68721: contig of 891 bp in length  
 33401 33502: gap of 100 bp \* 68721 68821: gap of 100 bp  
 33502 34376: contig of 875 bp in length \* 68821 69711: contig of 890 bp in length  
 34376 34377: gap of 100 bp \* 69711 69811: gap of 100 bp  
 34377 34477: contig of 879 bp in length \* 69811 70099: contig of 888 bp in length  
 34477 35355: contig of 100 bp \* 70099 70700: gap of 100 bp  
 35355 35455: gap of 100 bp \* 70700 71679: contig of 880 bp in length  
 35455 36370: contig of 915 bp in length \* 70800 70800: contig of 880 bp in length

Query Match 64.7%; Score 194; DB 2; Length 196472;  
 Best Local Similarity 79.3%; Pred. No. 5.3e+02;  
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CTGAGCCCTCTCGACTCGAGCTCCG 29  
 Db 6946 CTGGACCCCTCTGCTGAGATTC 6918

Search completed: March 1, 2005, 00:29:18  
 Job time : 883.065 secs

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: \_\_\_\_\_**

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)